

PRZYCZYNA I WYJAŚNIANIE

**STUDIUM Z FILOZOFII
I METODOLOGII NAUK**

PAWEŁ KAWALEC

mojej Żonie Annie oraz Stasiowi

Spis treści

Spis treści	5
Szczegółowy spis treści	7
Wykaz skrótów	11
Przedmowa	13
Wprowadzenie: Zarys dziejów kryterium przyczyny w filozofii i metodologii nauk	17

CZĘŚĆ I

PROCEDURALNE KRYTERIUM PRZYCZYNY

1. Nieeksperymentalne ustalenie przyczyny – studium przypadku	31
2. Proceduralne kryterium przyczyny	63
3. Kryterium proceduralne jako uogólnienie jakościowych kryteriów przyczyny	89
4. Typy ilościowych modeli przyczynowych	117

5. Kryterium proceduralne w odniesieniu do ilościowych modeli przyczynowych	195
---	-----

CZĘŚĆ II

ZARYS METODOLOGII PROCEDURALNEJ

6. Projekcyjność praw naukowych	215
7. Projekcyjność modeli przyczynowych	237
8. Zależności przyczynowe w wyjaśnianiu subsumpcyjnym	275
9. Projekcyjność w wyjaśnianiu przyczynowym	297
Zakończenie: Metodologia proceduralna – między realizmem a antyrealizmem naukowym	331
Dodatki	349
Bibliografia	403
Skorowidz terminów technicznych	435
Indeks	439

Szczegółowy spis treści

Spis treści	5
Szczegółowy spis treści	7
Wykaz skrótów	11
Przedmowa	13
Wprowadzenie: Zarys dziejów kryterium przyczyny w filozofii i metodologii nauk	17

CZĘŚĆ I

PROCEDURALNE KRYTERIUM PRZYCZYNY

1. Nieeksperymentalne ustalenie przyczyny – studium przypadku	31
1.1 Rys historyczny odkrycia przyczyny cholery przez J. Snowa	31
1.2 Dowody przenoszenia choroby w kontaktach bezpośrednich	38
1.3 Mechanizm rozwoju choroby w organizmie jednostki	41
1.4 Dowody przenoszenia choroby przez wodę	45
1.5 Interwencja a ustalanie przyczyny cholery	47
1.6 Eksperyment naturalny	50
1.7 Wyjaśnienie znanych zależności	58
1.8 Eliminacja hipotez alternatywnych	59
1.9 Odpowiedź na zarzuty	60

1.10	Profilaktyka jako test	60
2.	Proceduralne kryterium przyczyny	63
2.1	Warunki adekwatności kryterium przyczyny	64
2.2	Niezmienniczość przyczyny	73
2.3	Sformułowanie proceduralnego kryterium przyczyny	80
3.	Kryterium proceduralne jako uogólnienie jakościowych kryteriów przyczyny	89
3.1	B. Hilla kryteria przyczyny	89
3.2	M. Sussera kryteria przyczyny	94
3.3	K. Rothmana i S. Greenlanda krytyka kryteriów przyczyny	106
3.4	Jakościowe kryteria przyczyny a kryterium proceduralne	109
4.	Typy ilościowych modeli przyczynowych	117
4.1	Podziały modeli przyczynowych	118
4.2	Modele Neymana-Rubina (RCM)	127
4.3	Modele równań strukturalnych (SEM)	146
4.4	Modele sieci bayesowskich (BN)	169
5.	Kryterium proceduralne w odniesieniu do ilościowych modeli przyczynowych	195
5.1	Interpretacja przyczynowa modeli RCM wobec kryterium proceduralnego	196
5.2	Rozbieżności między typami modeli przyczynowych w świetle kryterium proceduralnego	200
5.3	Interpretacja przyczynowa modeli BN wobec kryterium proceduralnego	206

CZEŚĆ II

ZARYS METODOLOGII PROCEDURALNEJ

6.	Projekcyjność praw naukowych	215
6.1	Koncepcja projekcyjności uogólnień	217
6.2	Projekcyjność uogólnień a uniwersalność praw naukowych	227
6.3	Projekcyjność uogólnień a informatywność i prostota	230
6.4	Projekcyjność uogólnień a konieczność praw przyrody	233
7.	Projekcyjność modeli przyczynowych	237
7.1	Uogólnienia projekcyjne a regularnościowa koncepcja praw	239
7.2	Uogólnienia projekcyjne a koniecznościowa koncepcja praw	245

7.3	Uogólnienia projekcyjne a manipulacyjna koncepcja praw	251
7.4	Uogólnienia projekcyjne a relatywizm w kwestii praw	257
7.5	Uogólnienia projekcyjne a sceptycyzm w kwestii praw	263
8.	Zależności przyczynowe w wyjaśnianiu subsumpcyjnym	275
8.1	Zależności przyczynowe w wyjaśnianiu DN oraz IS	277
8.2	Zależności przyczynowe w wyjaśnianiu SR oraz CM	286
8.3	Zależności przyczynowe w wyjaśnianiu jako unifikacji	291
9.	Projekcyjność w wyjaśnianiu przyczynowym	297
9.1	Projekcyjność w manipulacyjnej koncepcji wyjaśniania	297
9.2	Projekcyjność w procesualnej koncepcji wyjaśniania	305
9.3	Wyjaśnianie przyczynowe jako subsumpcja proceduralna	312
9.4	Projekcyjność w proceduralnej koncepcji wyjaśniania	325
Zakończenie: Metodologia proceduralna – między realizmem a antyrealizmem naukowym		331
Dodatki		349
A:	Zmienne instrumentalne w modelu RCM	349
B:	S. Wrighta model przyczynowy SEM	361
C:	Autonomiczność relacji ekonomicznych – T. Haavelma model konsumpcji	365
D:	Przykład uaktualniania modelu przyczynowego BN	371
E:	Przykład modelu przyczynowego BN ze zmiennymi ukrytymi	379
F:	Modele przyczynowe BN jako uogólnienie modeli SEM	383
G:	Przykłady zastosowań modeli BN do ustalania przyczyn	393
Bibliografia		403
Skorowidz terminów technicznych		435
Indeks		439

Wykaz skrótów

- BN – sieć bayesowska (skrót ang. Bayesian network)
- CM – mechaniczyczne wyjaśnianie przyczynowe (skrót ang. causal mechanical)
- CSI – Komisja do Badań Naukowych (skrót ang. the Committee for Scientific Inquiries)
- DAG – acykliczny graf skierowany (skrót ang. directed acyclic graph)
- DN – wyjaśnianie dedukcyjno-nomologiczne (skrót ang. deductive-nomological)
- IS – wyjaśnianie indukcyjno-statystyczne (skrót ang. inductive-statistical)
- ITT – leczenie zamierzone (skrót ang. intention-to-treat)
- MLE – maksymalny estymator wiarygodności (skrót ang. Maximum Likelihood Estimator)
- RCM – Rubina model przyczynowy (skrót ang. Rubin Causal Model)
- SEM – model równań strukturalnych (skrót ang. structural equation model)
- SR – relewancyjne wyjaśnianie statystyczne (skrót ang. statistical relevance)
- SUTVA – założenie o stabilności oddziaływania na jednostkę (skrót ang. Stable Unit Treatment Value Assumption)

Przedmowa

Problematyka związana z zależnościami przyczynowymi, ich modelowaniem i odkrywaniem, po długiej nieobecności w filozofii i metodologii nauk, budzi współcześnie duże zainteresowanie. Wiąże się to przede wszystkim z dynamicznym rozwojem, zwłaszcza od lat 1990., technik obliczeniowych. Wypracowane w tym czasie sieci bayesowskie uznaje się za matematyczny język przyczynowości. Pozwalają one na daleko idącą automatyzację wnioskowań, co jest także zachętą do podjęcia prób algorytmizacji odkrywania przyczyn.

Na potrzeby badań naukowych, które pozwalają na przeprowadzenie eksperymentu z randomizacją, standardowe metody ustalania zależności przyczynowych opracowano na początku XX wieku. Zupełnie inaczej sprawa przedstawia się w przypadku badań nieeksperymentalnych, gdzie podobne rozwiązania pozostają kwestią przyszłości.

Zadaniem tej książki jest podanie warunków, które powinny być spełnione przez te rozwiązania, oraz sformułowanie proceduralnego kryterium zależności przyczynowych jako szczegółowej realizacji tych warunków. Pociąga ono ważne konsekwencje dla filozofii i metodologii nauk, które ujawnia – podany w Części II – zarys metodologii proceduralnej.

W literaturze przedmiotu brakuje w miarę wszechstronnego i systematycznego omówienia najnowszych filozoficznych i metodologicznych dyskusji na temat przyczynowości, co niech będzie wytłumaczeniem, dlaczego w niektórych punktach obecnej książki szczegółowo referuję trudno dostępne teksty źródłowe.

Przymiotnik „proceduralny” używam tu w znaczeniu węższym niż Huw Price (w którego pracach właściwszy byłby termin „kryterialny”) dla

podkreślenia – zgodnie z łacińskim źródłosłowem *procedo* – że dla ustalenia przyczyny niezbędne jest podjęcie przez uczonych określonych *interakcji* z badaną rzeczywistością.

Załączki zamysłu prezentowanego w tej książce przedstawiłem podczas warsztatów filozoficznych „Philosophy and Probability” w roku 2002, zorganizowanych przez Instytut Filozofii Uniwersytetu w Konstancji. Wdzięczny jestem uczestnikom tych warsztatów za uwagi, a przede wszystkim następującym osobom: Lucowi Bovensowi, Brandonowi Fitelsonowi, Alanowi Hájekowi, Stephanowi Hartmannowi oraz Jonowi Williamsonowi. Podczas międzynarodowej konferencji „Analytical Pragmatism”, zorganizowanej w Lublinie w roku 2003 przez Wydział Filozofii Katolickiego Uniwersytetu Lubelskiego, odniosłem swoją koncepcję do prac Nancy Cartwright. Szczególnie inspirujący okazał się komentarz Huw Price’a do mojego referatu i przeprowadzona z nim dyskusja.

Ujęcie koncepcji metodologii proceduralnej na szerszym tle współczesnego nurtu empirystycznego w filozofii nauki przedstawiłem w roku 2004 podczas konferencji „5th Quadrennial Fellows Conference”, zorganizowanej przez Instytut Filozofii Uniwersytetu Jagiellońskiego oraz Centrum Filozofii Nauki w Pittsburghu. Szczególnie pomocne w dalszych moich pracach były uwagi Jamesa Bogaena, Janet Kourany, Jamesa Lennox, Johna Nortona, Thomasa Bonka, Jana Woleńskiego i Johna Worralla, za które wyrażam swoją wdzięczność.

Korpus książki powstał podczas mojego stażu w Centrum Filozofii Nauki w Pittsburghu, który odbyłem jako stypendysta Fundacji na Rzecz Nauki Polskiej w roku akademickim 2004-2005. Uczestniczyłem w tym czasie w życiu naukowym Centrum i w pracach badawczych zespołu z Instytutu Filozofii Uniwersytetu Carnegie-Mellon w Pittsburghu kierowanego przez Clarka Glymoura. Na jego ręce składam podziękowanie za wiele pomocnych uwag do moich wystąpień oraz tekstów i za dyskusje przede wszystkim z nim samym i z jego najbliższymi współpracownikami: Peterem Spirtesem oraz Richardem Scheinesem, a także pozostałymi członkami tego zespołu, doktorantami i uczestnikami seminarium badawczego „Causality in the Social Sciences”.

Za wieloletnie wsparcie, wielopłaszczyznowe inspiracje towarzyszące pisaniu tej książki, a także liczne pomocne uwagi do jej wcześniejszych wersji dziękuję przede wszystkim Księdzu Profesorowi Andrzejowi Bronkowi oraz Księdzu Profesorowi Józefowi Herbutowi, współprowadzącemu seminarium

doktorskie w Katedrze Metodologii Nauk Katolickiego Uniwersytetu Lubelskiego Jana Pawła II, jak również pozostałym uczestnikom tego seminarium.

Dziękuję mojej Żonie, dr Annie Kawalec, za wiele wysiłku włożonego w ulepszenie redakcji – językowej i merytorycznej – obecnej książki.

Książkę tę można czytać na kilka sposobów. Czytelnikom zainteresowanym przede wszystkim prowadzeniem badań empirycznych polecałbym rozpoczęcie od Rozdziału 2. i kontynuację pozostałych rozdziałów Części I, a następnie Dodatków.

Czytelnikom zainteresowanym problemami filozofii i metodologii nauk polecałbym rozpoczęcie lektury książki od Części II i uzupełniającą lekturę Rozdziału 2., a następnie Wprowadzenia i Zakończenia.

Czytelnikom mniej zainteresowanym zagadnieniami teoretycznymi polecałbym zapoznanie się z fascynującymi dziejami odkrycia przyczyn cholery przez Johna Snowa, które rekonstruuje w Rozdziale 1. W dalszej części natomiast polecałbym przejście do Wprowadzenia i Zakończenia, które w mniej specjalistyczny sposób przybliżają proponowane tu rozstrzygnięcia.

Tekst książki nie był dotąd publikowany. Wyjątkiem są pewne fragmenty Rozdziału 8. oraz 9., które w zmienionej postaci ukazały się w *Rocznikach Filozoficznych* (Kawalec 2004).

Lublin, luty 2006 r.

Wprowadzenie

Zarys dziejów kryterium przyczyny w filozofii i metodologii nauk

Naukowe poznanie świata w jego bogactwie, różnorodności, ale i złożoności nieodłącznie związane jest z ustalaniem przyczyn. Pozwalają one na wyjaśnienie zaobserwowanego stanu rzeczy. Nic dziwnego więc w tym, że od zarania refleksji nad nauką stale towarzyszą jej te dwa kluczowe tematy, którymi są przyczynowość i wyjaśnianie.

Kilkanaście stuleci poświęconych im rozważań zrodziło bardzo złożoną problematykę i ogromną literaturę. Próba podania dziś wyczerpującego ujęcia obu tych zagadnień wydaje mi się przedsięwzięciem zbyt śmiałym. Owszem, można nie dostrzegać złożoności problematyki i omijać trudniejsze tematy; można jednak ograniczyć się do pewnego obszaru, by na nim skupić swoją uwagę. W tej książce przyjmuję to drugie rozwiązanie i dlatego rozpocznę od zarysowania granic, w obrębie których pozostaje niniejsze studium.

Kilkuletnie badania nad przyczynowością i wyjaśnianiem oraz zgłębianie literatury specjalistycznej potwierdziły moje przeświadczenie o tym, że współcześnie w badaniu przyczynowości wyodrębniają się dwa, stosunkowo niezależne, obszary. Jeden z nich to *filozoficzne* zagadnienia związane z naturą przyczynowości, skoncentrowane przede wszystkim na pytaniach: co to jest przyczynowość oraz czym jest przyczynowość. Za R. Ingardenem chciałbym podkreślić, że charakter odpowiedzi na te pytania, a co za tym idzie – filozoficzna koncepcja przyczynowości – uzależnione są w pierwszym rzędzie od tego, jak wyartykułowana zostanie konieczność związku

przyczynowego.¹ Zagadnienia pokrewne dotyczą tego, jakie są rodzaje przyczynowości² i jakie zachodzą między nimi związki.³ Kontynuacją tej problematyki są obszerne tematy takie, jak zasada przyczynowości czy determinizm przyczynowy.⁴

Drugie pole badań skupia się wokół wskazania *kryteriów* przyczynowości. Kryteria te mają pozwolić na stwierdzenie faktycznego zachodzenia zależności przyczynowej między badanymi zjawiskami, a podstawą tego ustalenia są dane, które pozyskano w badaniach naukowych. Czy spożywanie alkoholu powoduje choroby serca? Czy zwiększenie inflacji wpływa na obniżenie bezrobocia? Czy zaostrenie kar, a w szczególności wprowadzenie kary śmierci, zmniejsza przestępczość? Są to przykłady pytań wymagające wyraźnego określenia kryterium przyczynowości, gdyż trudno jest w inny sposób znaleźć na nie odpowiedź. Zasadniczym utrudnieniem jest niemożność posłużenia się eksperymentem, który pozwoliłby precyzyjnie ocenić siłę wpływu jednego tylko czynnika, minimalizując wpływ pozostałych.⁵ Jak krótko wyraził to A. Sulek:

Zwykle jednak eksperyment służy sprawdzeniu hipotez o związkach przyczynowych między czynnikami zmienianymi a pewnymi innymi. Drogą do tego celu wiodącą jest kontrolowane wywoływanie zmiany, której skutków badacz jest ciekawy. (1979, 101)

Innym powodem trudności w eksperymentalnym ustalaniu przyczyny może być upływ czasu, który niekiedy jest konieczny do wychwycenia danej zależności, np. gdy chcemy ocenić szkodliwość akumulacji ołowiu wdychanego z zanieczyszczonego spalinami powietrza. Jeszcze innym powodem jest niekontrolowany wpływ kilku czynników, które zakłócają badaną zależ-

¹ (Ingarden 1945, 151-61). Sformułowanie Ingardena jest, jak sądzę, szerokie i obejmuje anomalizm D. Davidsona (1970/1992; por. też Żegleń 2003, 68-76), jak również stanowisko G. E. M. Anscombe, zastępującej konieczność (w sensie wąskim) przez wyprowadzalność (Anscombe 1971/1981, 136).

² Pluralizmu pojęć przyczyny broni C. Hitchcock (2003a; 2003b).

³ Autorzy związani z London School of Economics, zwłaszcza N. Cartwright i J. Williamson, pod wpływem K. R. Poppera (1990/1996, 11-37) łączą filozoficzną analizę pojęcia przyczynowości z – przede wszystkim skłonnościową – interpretacją pojęcia prawdopodobieństwa.

⁴ Por. (Metallmann 1934, 265-373; Bunge 1959/1968; Krajewski 1967, 241-59; Gawęcki 1969, 22-82; Heller 2002, 23-25).

⁵ Przystępny zarys zagadnień dotyczących badania zachodzenia zależności przyczynowych w badaniach eksperymentalnych w literaturze przedmiotu podał J. Karpiński (1985, 72-88) i E. Edelman (2001).

ność przyczynową, np. przy ocenie skuteczności zaostrzania kar w zapobieganiu przestępczości.

Podzielałam zdanie autorów, którzy dostrzegając związek filozoficznego i kryterialnego podejścia⁶ do badań nad przyczynowością, podkreślają ich zasadniczą odrębność.⁷ Zwięzłe różnicę tę wysławiają Peter Menzies i Huw Price:⁸

odrzucaamy każdy pogląd, który stwierdza, że znaczenie twierdzeń przyczynowych ma być określone w postaci ich warunków weryfikacji lub uzasadnionej asercji. Naszym zdaniem pojęcie sprawczości występuje w analizie przyczynowości nie z tego powodu, że eksperymentalne manipulacje dostarczają dowodów twierdzeniem przyczynowym, ale ponieważ pojęcie przyczynowości jest pojęciem, które ze swojej natury jest zakorzenione w idei manipulacji. (Menzies i Price 1993, 196)⁹

Uwzględniając stosunkową odrębność problematyki obu obszarów dociekań nad przyczynowością, zdecydowałam, by zamiast uproszczeń i pobieżnej charakterystyki stanowisk filozoficznych w kwestii natury przyczynowości, zostawić tę problematykę poza ramami obecnej książki. Zasadniczym więc tematem Części I jest dyskusja nad różnymi kryteriami przyczynowości i moja próba ich uogólnienia w postaci *kryterium proceduralnego*. Ponieważ są to zagadnienia szczególnie trudne w przypadku badań nieeksperymentalnych, dlatego też im przede wszystkim poświęcam swoją uwagę.¹⁰

Zgodnie z koncepcją proceduralną charakteryzowanie związków przyczynowych¹¹ prowadzi, jak staram się pokazać w Części II, do istotnych zmian w określeniu tego, na czym polega podanie wyjaśnienia. Wyjaśnianie jako *subsumpcja proceduralna* integruje dwa rozdzielane dotąd rodzaje wyjaśniania: jako podania przyczyn oraz jako podporządkowania ogólnej prawidłowości. W proponowanej tu koncepcji podanie wyjaśnienia nie jest

⁶ W. Krajewski proponował określać tego rodzaju podejście jako „kondycjonalizm” (1967, 86-105).

⁷ Na powszechny brak odróżniania zagadnień filozoficznych od zagadnień kryterialnych w filozofii współczesnej sugestywnie zwraca uwagę S. Judycki (2005).

⁸ Tę różnicę respektowali klasycy przyczynowości, a w szczególności – jak pokazuje J. Mackie (1974, 90) – D. Hume i I. Kant. Jej zastosowanie we współczesnych dyskusjach o przyczynowości mentalnej dyskutuje U. Żegleń (2003, 174).

⁹ W przypadku tekstów obcojęzycznych przywoływanych w tej książce tłumaczenia dokonał P. Kawalec.

¹⁰ Podstawowe zarzuty przeciw zrandomizowanym eksperymentom zwięzłe prezentują C. Howson i P. Urbach (1993, 374-82). Por. też (Ferguson i Takane 1989/2003, 266-67).

¹¹ Ich typologię systematycznie dyskutował Krajewski (1967, 116-72).

uzależnione, jak przyjmowano tradycyjnie, od podania ścisłych praw. W tym więc zakresie odrzucone mogą być niektóre sztuczne problemy w metodologii nauk społecznych, mające źródło w mechanicznym ich zestawianiu z naukami przyrodniczymi.¹²

*

Szczegółową prezentację zagadnień poruszanych w książce poprzedzę zarysem dziejów pojęcia przyczynowości w ogólnej metodologii i filozofii nauki. Jak często obserwujemy w przypadku dziejów idei, dzieje tego pojęcia przebiegały falami. W starożytności, średniowieczu i nowożytności traktowano pojęcie przyczynowości jako kluczowy składnik metod naukowych i rozumienia nauki.¹³ Krytyka D. Hume'a podważyła to przeświadczenie i pośrednio wpłynęła na wyrugowanie jawnego odwoływania się do pojęcia przyczynowości we współcześnie stosowanej statystycznej analizie danych. Od lat 1980. dokonała się jednak znacząca zmiana w kierunku przywrócenia wcześniej przypisywanej pojęciu przyczynowości roli w nauce. Koło, jakie zatoczyła w tym przypadku historia nauki, nie jest błędne – współcześnie pojęciu przyczynowości towarzyszą bardzo precyzyjnie określone kryteria jego stosowania.

Arystoteles, charakteryzując naturę wiedzy naukowej, odwołał się do pojęcia przyczynowości, łącząc je w ten sposób trwale przede wszystkim z zagadnieniem wyjaśniania naukowego:

Sądzimy, że wówczas poznajemy coś bezwarunkowo, a nie w sposób sofistyczny, akcydentalny, gdy jesteśmy przekonani, że poznaliśmy przyczynę, dzięki której rzecz istnieje, że jest rzeczywistą przyczyną, i że inaczej być nie może. (Arystoteles *Analityki wtóre*, Ks. I, R. 2; 1990, 256)

Był on przede wszystkim zainteresowany wyartykułowaniem filozoficznej teorii przyczynowości i jej rodzajów. Najbardziej drobiazgowo dopracowana przez Arystotelesa teoria, którą była sylogistyka, jak podkreśla T. Czeżowski (1933, 39), podawała „metodę wiązania wiedzy o rzeczach z wiedzą o przyczynach”. I trudno byłoby uznać przeprowadzoną przez Arystotelesa

¹² Por. (Mokrzycki 1980, 159-60).

¹³ Zarys historyczny podają m.in. T. Czeżowski (1933), J. Metallmann (1934, 123-30); M. Bunge (1959/1968, 76-114) i S. Kiczuk (1995, 11-27).

analizę przyczynowości za zmierzającą do ustalenia kryteriów odróżniających zależności przyczynowe od przypadkowych. J. Losee twierdzi, że takim kryterium według Arystotelesa miało być określenie rodzaju atrybutu, który w przypadku relacji przyczynowej:

(1) jest prawdziwy w stosunku do każdego egzemplarza podmiotu, (2) jest prawdziwy w odniesieniu do podmiotu jako takiego, a nie jako części większej całości, (3) jest „istotny” dla tego podmiotu. (Losee 1993/2001, 20)

Nie wdając się w szczegółową dyskusję, zwrócę tylko uwagę, że analizując te rzekome kryteria, Losee uznaje je za niewystarczające do wskazania zależności przyczynowych w wiedzy faktualnej.¹⁴ Potwierdza to, moim zdaniem, słuszność tezy, że w tym przypadku Arystoteles realizował inne zadanie, jakim było opracowanie filozoficznej koncepcji natury przyczynowości.¹⁵

Arystoteles, wiążąc zagadnienia przyczynowości i wyjaśniania, wywarł w tym względzie znaczący wpływ na późniejszych teoretyków nauki. Konieczność związku przyczynowego, podobnie jak Arystoteles, przenosili oni jednak na wyjaśnianie naukowe. Było to poważne ograniczenie w rozumieniu nauki, nawet w odniesieniu do ówczesnego stanu wiedzy:

Arystotelesa koncepcja nauki stawiała jej bardzo rygorystyczne wymagania. Postulowała, aby z rozważania zjawisk wysnuć pewne i ogólne pryncypia o charakterze określeń, z których dedukuje się prawdy przyczynowo wyjaśniające naturę rzeczy jedynie w celach teoretycznych. Nawet filozofia pierwsza, która miała być naukową wiedzą przede wszystkim, nie spełniała faktycznie tych wszystkich warunków. Tym bardziej nie spełniały ich dyscypliny przyrodnicze [...]. (Kamiński 1992, 60)

W okresie tzw. rewolucji naukowej w nowożytności, mimo wielu zmian, które dokonały się w rozumieniu nauki i jej uprawianiu, wprowadzony przez Arystotelesa związek między przyczynowością a wyjaśnianiem w zasadzie nie podlegał kwestionowaniu. Ilustracji tego dostarczają podane przez I. Newtona „reguły rozumowania w filozofii”, które miały określać sposób znajdowania ogólnych prawdowości o funkcji eksplanacyjnej. Szczególnie istotne są dwie:

¹⁴ Odróżniana przez Arystotelesa od eksplanacyjnej wiedzy faktualną charakteryzuje bliżej S. Kamiński (1992, 54).

¹⁵ O Arystotelesa koncepcji wiedzy naukowej szerzej pisze T. Kwiatkowski (1969).

Reguła I

Nie powinniśmy przyjmować więcej przyczyn dla przedmiotów świata przyrodniczego niż te, które są zarazem prawdziwe i wystarczające do wyjaśnienia tego, jak one nam się jawią. [...]

Reguła II

Zatem, tym samym skutkom w świecie przyrody musimy, na ile to tylko możliwe, przypisywać te same przyczyny. (Newton 1687/1964, 324)

Mimo wyraźnego podkreślenia różnicy między „przedmiotami świata przyrodniczego” a „tym, jak one nam się jawią”, czyli badanymi w nauce zjawiskami, rola przyczynowości w wyjaśnianiu tychże zjawisk nie podlegała dla Newtona wątpliwości.

Utrzymywanie zachodzenia związku między uogólnieniami przyczynowymi a wyjaśnianiem nie było dla Newtona przeszkodą w odrzuceniu koniecznego charakteru poznania świata przyrody. Nie można stwierdzić z całą pewnością, że regularności rządzące zjawiskami nie mogłyby być inne, a więc istnieje możliwość, że w świetle nowych danych dotychczasowe uogólnienia zostaną zrewidowane.

Funkcję eksplanacyjną uogólnień przyczynowych radykalnie zaatakował dopiero Hume przy okazji krytyki koniecznego charakteru wiedzy naukowej.¹⁶ Zaobserwowana regularność występowania zjawiska B po zjawisku A nie jest wystarczającą podstawą do stwierdzenia konieczności tego związku, którą można by wyrazić przez warunek kontrfaktyczny: w przyszłości nie jest możliwe wystąpienie A, któremu nie towarzyszyłoby B.

Zwróćmy jednak uwagę, że sceptycyzm przyczynowy Hume’a zorientowany był wyłącznie przeciw zadaniu, które wcześniej określiłem jako sformułowanie filozoficznej teorii przyczynowości, ale nie przeciw podaniu kryteriów przyczynowości. *Badania* – w stosunku do wcześniejszego *Traktatu*, gdzie przyczynowość określona była jako regularność wystąpień A i B – uzupełnił on przez dodanie warunku kontrfaktycznego (Hume 1748/1977, 93), a w *Traktacie* umieścił listę reguł „wskazujących, jak sądzić o przyczynach i skutkach” (Hume 1739/1963, 227).¹⁷

Odpowiedzią na sceptycyzm przyczynowy Hume’a oraz wynikające z niego podważenie zasadności wnioskowań indukcyjnych była matematyczna

¹⁶ Analizę stanowiska Hume’a przeprowadza Mackie (1974, 3-28).

¹⁷ Szczegółową analizę podanych przez Hume’a trzech definicji przyczynowości w literaturze polskiej przeprowadził A. Wawrzyniak (1964, 44-51).

teoria wnioskowań indukcyjnych Thomasa Bayesa.¹⁸ Stanowiła ona popularny w XIX wieku poprzednik instensywnie rozwijanych od lat 1990. nowoczesnych metod statystycznej analizy danych.

Można postawić pytanie, dlaczego w okresie drugiej połowy XIX wieku i pierwszych dekadach XX wieku, które miały decydujący wpływ na ukształtowanie się współczesnej statystyki i metod analizy danych, do podręczników trafiły nie metody bayesowskie właśnie, lecz tzw. klasyczne metody wnioskowań statystycznych, które rugowały przyczynowość poza naukę. Jest to pytanie intrygujące, gdyż powstające wówczas modele zależności zjawisk, np. analizy ścieżkowej¹⁹ Sewalla Wrighta, były z zamierzenia modelami²⁰ przyczynowymi.

Z pewnością można podać wiele wiarygodnych odpowiedzi tłumaczących rozszerzenie filozoficznego sceptycyzmu przyczynowego Hume'a na zagadnienie kryteriów zależności przyczynowych, które dokonało się na przełomie XIX i XX wieku. Jego autorami było niewątpliwie grono wybitnych myślicieli z uniwersytetu w Cambridge takich, jak R. A. Fisher, K. Pearson, J. Neyman, E. Pearson, H. Jeffreys, J. Keynes i B. Russell.

S. Kamiński, śledząc dzieje nauki i pojęcia nauki, zwraca uwagę na powszechnie – za A. Comte'em i E. Machem – przyjmowaną wówczas teorię nauki:

Chodzi o stwierdzenie i czyste opisanie faktów pod względem ich korelacji i sekwencji, nie zaś o ich wyjaśnienie. Nauka mówi „że” i „jak” zachodzą zdarzenia, natomiast nie jest zdolna odpowiedzieć na pytanie „dlaczego?” [...]. Nauki przyrodnicze nie osiągnęły jeszcze stanu doskonałego. Zawierają wiele pojęć niezgodnych z naturą, jak np. przyczynowość, materia, siła. Te pojęcia metafizyczne trzeba eliminować. (Kamiński 1992, 111)

W tradycyjnej empirystycznej metodologii, która dominowała w logicznym empiryzmie, a pod jego wpływem w myśli pierwszej połowy XX wieku – jej ślady ujawniają się niekiedy i dziś jeszcze – zależności przyczynowe rozumiano redukcyjnie.²¹ Stwierdzenie, że między dwoma zjawiskami

¹⁸ Rozwinięcie i uzasadnienie tej hipotezy za S. Zabellem (1997) podaje Kawalec (2003b).

¹⁹ Por. m.in. (Gaul i Machowski A. 1987).

²⁰ Systematyczną dyskusję różnych pojęć modelu występujących w naukach empirycznych podaje R. Wójcicki (1979). Z wyróżnionych przez niego pojęć modelu, przyjęte w tej pracy najbliższe jest kategorii „obiektów matematycznych jako modeli operacyjnych” (Wójcicki 1974, 290-294). Por. też (Mejbaum 1995, 146-149). Filozoficzne zagadnienia teorii modelowania dyskutuje szerzej M. Lubański (1997, 109-16).

²¹ Stanowisko to wyartykułował m.in. M. Schlick (1932/1949).

zachodzi zależność przyczynowa, zredukowano do stwierdzenia regularności współwystępowania tych zjawisk oraz do asymetrii czasowej między nimi. Tym samym więc rugowano zależności przyczynowe jako wtórne i w zasadzie zbędne wobec możliwości, jakie daje wnioskowanie statystyczne,²² co T. Czeżowski przypisał „niejasności związku przyczynowego” (1946, 114).

Jedną z konsekwencji takiego stanowiska w tradycyjnej metodologii są – na co zwracam uwagę w Części II – problemy związane z pojęciem prawa naukowego. Z jednej strony bywa ono przeceniane jako podstawa wszelkiego wyjaśniania naukowego. Z drugiej zaś – zwłaszcza na gruncie współczesnego empiryzmu (van Fraassen 1989) – ogranicza się wartość poznawczą praw i zakres ich obowiązywalności, a czasem całkowicie ruguje się je poza naukę, jako filozoficzne artefakty.

W latach 1930. zaczęły jednak powstawać podwaliny teoretyczne zmiany, jaka współcześnie dokonuje się w badaniach wykorzystujących metody statystyczne. Wyrugowane poza naukę pojęcie przyczynowości ma współcześnie precyzyjnie określone kryteria stosowane nie tylko do tworzenia modeli przyczynowych, ale również do odkrywania zależności przyczynowych w danych. Prace F. P. Ramseya oraz B. de Finettiego ugruntowały podstawy teoretyczne statystyki bayesowskiej.²³

Niemal równolegle realizowane były inne przedsięwzięcia, które łącznie złożyły się na zintegrowaną współcześnie w graficznych modelach przyczynowych reprezentację i metody odkrywania zależności przyczynowych. Zwróćmy tu uwagę na dwa szczególnie istotne przedsięwzięcia. Jednym były prace H. Reichenbacha nad definicją relewancji przyczynowej w kategoriach probabilistycznych, a zwłaszcza podanie ścisłej definicji pojęcia wspólnej przyczyny, dzięki któremu można byłoby precyzyjnie odróżnić

²² Współczesną i bardziej subtelną wersję redukcjonizmu proponuje Karhausen (2001, 66): „W naukach stosowanych, medycynie i epidemiologii, przyczyny ujawniają, co należałoby zrobić, aby osiągnąć określony rodzaj wyników; są nieodłącznie związane z celami i skutecznymi strategiami; są zmianami (faktami, zdarzeniami czy czynami), które mogą być wykorzystane w sposób szkodliwy lub udany; są kontrastowe, ponieważ wnoszą różnicę do okoliczności, w których występują, oraz do tych, w których nie występują. Każda definicja przyczynowości nieuchronnie sprowadzi się do użycia metod epidemiologicznych. Postępujące metodologiczne wyrafinowanie ostatnich czterdziestu lat jest w doskonałej zgodności ze stopniową niejawną wyprzedząją pojęcia przyczynowości.”

²³ Por. też (Corfield i Williamson 2001; Dale 1991; Earman 1992; Fisher 1973, 8-39; Gillies 2000; Howie 2002; Howson i Urbach 1993, 169-441; Osiewalski 2001, 8-9; Seidenfeld 1979, 1-102; Swinburne 2002; Szaniawski 1961/1994, 197-207; Szaniawski 1967/1994; Woleński 2001, 150-52).

faktycznie zachodzącą zależność przyczynową od pozornej.²⁴ Stąd też pojęcie wspólnej przyczyny stało się jednym z kluczowych pojęć we współczesnych modelach przyczynowych.²⁵ Wypracowana przy tej okazji przez Reichenbacha interpretacja pojęcia prawdopodobieństwa²⁶ – jako granicy częstości względnej – dała asumpt do wykształcenia się – głównie za pośrednictwem prac H. Putnama – formalnej teorii uczenia się, którą zintegrowano w zautomatyzowanych metodach odkrycia przyczynowego.²⁷

Drugim ważnym nurtem było zainicjowane przez działającą w latach 1940. Komisję Cowlesa wypracowanie metod modelowania ekonometrycznego, które pozwalałyby nie tylko na reprezentację zaobserwowanych korelacji, ale przede wszystkim na przewidywanie dzięki tym modelom efektów podejmowanych działań. Rezultatem tych prac było wypracowanie koncepcji modeli równań strukturalnych, w których każdy współczynnik wyraża wielkość odpowiadającą zmianie, jaka miałaby miejsce, gdyby pozostałe przyczyny nie ulegały zmianie. Judea Pearl stopniowy zanik jawnej interpretacji przyczynowej modeli równań strukturalnych przypisuje przede wszystkim temu, że twórcy tej koncepcji, jak T. Koopmans czy T. Haavelmo, nie wypracowali adekwatnej notacji, oddającej asymetrię zależności przyczynowej.

W latach 1980. dokonał się zasadniczy przełom w ignorowaniu – często pozornym, jak pokazuje choćby przykład bardzo rozpowszechnionych modeli równań strukturalnych – przyczynowości w analizie danych. Sądzę, że składają się na niego trzy poniekąd niezależne dokonania. Z pewnością decydujące znaczenie miało sformułowanie przez T. Speeda i H. Kiiveri (1982) tzw. założenia Markowa, który obecnie jest podstawą przyczynowej interpretacji modeli graficznych, oraz wykazanie, że jest on uogólnieniem wszystkich stosowanych w badaniach metod analizy zależności zmiennych.

²⁴ R. Wójcicki proponuje inne rozwiązanie: „Odkrycie zależności przyczynowej wymaga czegoś więcej niż stwierdzenia pozytywnej korelacji. To prawa wyznaczające mechanizm zdarzeń wyznaczają związki przyczynowe. Niekoniecznie natomiast istnienie pozytywnej korelacji między pewnymi zdarzeniami stanowi wyraz bezpośrednich powiązań przyczynowych między nimi” (1995, 34).

²⁵ Por. (Kamlah 1991; Hofer-Szabó i in. 1999).

²⁶ Przystępne wprowadzenie do tej problematyki opracowała H. Mortimer (1982, 20-29).

²⁷ Więcej szczegółów historycznych podaje Kawalec (2005). Początkową fazę prac nad automatyzacją odkryć przyczynowych źródłowo referują P. Langley i in. (1987, 302-304). Problemy współczesne omawiają m.in. Hendry i Krolzig 2004; Granger i Hendry 2005; Hansen 2005; Hoover 2005; Kuersteiner 2005; Leeb i Pötscher 2005; Paruolo 2005; Phillips 2005a; 2005b; 2005c).

Zasadniczym walorem zależności przyczynowych w stosunku do czysto statystycznych jest ich „stabilność”. Zależności poprawnie rozpoznane na podstawie wcześniej zebranego materiału empirycznego jako przyczynowe pozostają bez zmian nie tylko w stosunku do danych uzyskanych w kolejnych obserwacjach, ale także w stosunku do danych pochodzących z oddziaływania przez człowieka na badany wycinek rzeczywistości.

Drugim decydującym krokiem było wypracowanie przez Pearla architektury modeli graficznych w postaci tzw. sieci bayesowskich, które są współcześnie szeroko stosowane w programowaniu. Z uwagi na asymetrię występujących w nich zależności oraz łatwość ich zautomatyzowanego uaktualniania w oparciu o napływające dane, są one szczególnie cenione w konstruowaniu modeli przyczynowych. S. Lauritzen (2001) zmodyfikował przyjmowaną przez Pearla notację, dopełniając tym samym krok, którego w latach 1940. nie wykonali twórcy modeli równań strukturalnych.

C. Glymour i jego współpracownicy sformułowali dodatkowe założenia, których wprowadzenie do graficznych modeli przyczynowych pozwala na automatyzację procesu odkrywania zależności przyczynowych na podstawie danych.²⁸ W społeczności statystyków wprowadzenie algorytmów odkrycia przyczynowego budzi jednak wiele wątpliwości.

*

W literaturze przedmiotu panuje rozbieżność co do tego, jaką postać powinno mieć kryterium przyczynowości w przypadku badań nieeksperymentalnych i dlatego też w pierwszej części swoich rozważań poddałem systematycznemu studium zagadnienie kryterium zależności przyczynowej. W Rozdziale 1. przywołałem źródłowo najważniejszy z historycznie znanych przypadków odkrycia zależności przyczynowych, za jaki powszechnie uznaje się odkrycie przez angielskiego lekarza J. Snowa przyczyn epidemii cholery, które trzykrotnie nawiedzały XIX-wieczny Londyn.

Dzięki wyodrębnieniu najistotniejszych ogniw argumentacji i postępowania badawczego Snowa opracowałem własną propozycję kryterium zależności przyczynowej, które określam jako kryterium *proceduralne*. Zasada się ono na, moim zdaniem, pierwotnym w metodologii nauk charakte-

²⁸ O zaletach automatyzacji odkryć, nie tylko przyczynowych, w nauce pisze szerzej J. Żytkow (1995, 57-58).

rze pojęcia interakcji badawczej z rzeczywistością, określa też warunki, jakie muszą być spełnione, by dane, poddane obróbce statystycznej, mogły być podstawą uznania *faktycznego* zachodzenia zależności przyczynowej. Precyzyjne sformułowanie proceduralnego kryterium przyczynowości podają w Rozdziale 2.

W pozostałych rozdziałach Części I pokazują, że kryterium proceduralne jest uogólnieniem kryteriów zależności przyczynowych dotychczas proponowanych w literaturze metodologicznej. W perspektywie kryterium proceduralnego w Rozdziale 3. podają analizie nieformalne kryteria szeroko rozpowszechnione w badaniach nieeksperymentalnych, a w Rozdziale 5. formalne kryteria stosowane w podstawowych rodzajach modeli zależności przyczynowych, jakie stosuje się współcześnie w badaniach nieeksperymentalnych. Wyróżnienie rodzajów tych modeli i krótką ich charakterystykę podają w Rozdziale 4., a w Dodatkach szczegółowiej prezentują pewne zagadnienia związane z modelami graficznymi oraz dokonują porównania poszczególnych rodzajów modeli.

Sednem „cichej rewolucji”, jaka rozpoczęła się w końcu lat 1980.,²⁹ jest – według mnie – zmiana przyjętego dość powszechnie rozumienia nauki i coraz dalej sięgająca automatyzacja badań naukowych (por. Glymour 2004a). Nie sposób w tej pracy wyłożyć tego zagadnienia w sposób pełny (por. Kawalec 2005), dlatego ograniczam się tutaj do zasadniczych dla rozumienia nauki kategorii, jakimi są prawo naukowe oraz wyjaśnianie. Temu tematowi poświęcam Część II książki.

Jedną ze składowych trwałej spuścizny logicznego empiryzmu we współczesnej refleksji nad nauką jest uzależnianie naukowości badań od możliwości sformułowania na ich podstawie rygorystycznie pojmowanego prawa naukowego. Dlatego też w Rozdziale 6. proponuję zastąpienie pojęcia prawa naukowego przez pojęcie uogólnienia projekcyjnego, które – jak szczegółowo argumentuję – posiada wszystkie najistotniejsze walory praw, ale nie narzuca zbyt restryktywnych ograniczeń na badania naukowe. Szczegółową charakterystykę projekcyjności uogólnień modeli przyczynowych formułuję w Rozdziale 7. Zasadniczy problem, jaki wymaga tutaj rozstrzygnięcia, to rozbieżność, która wynika z projekcyjności uogólnień, a z drugiej strony, z proceduralności modeli przyczynowych.

²⁹ Zasadniczą dla wprowadzonej przez Pearla i zestandaryzowanej przez Lauritzena notacji ideę traktowania przyczynowości w perspektywie interwencji sformułował znacznie wcześniej G. von Wright (1971).

W Rozdziałach 8. i 9. pokazuję niedostatki różnych koncepcji wyjaśniania naukowego. Zasadniczym, według mnie, tego powodem jest nieodróżnianie dwóch rodzajów podejść do zagadnienia przyczynowości. Jednym z nich jest podejście filozoficzne, w którym zasadnicze pytanie dotyczy natury przyczynowości, realnej lub konceptualnej. Drugim rodzajem jest podejście, które określam tu jako kryterialne, gdzie kwestią zasadniczą jest ustalenie kryteriów zależności przyczynowych. Oba rodzaje podejść są ze sobą związane, jednak zagadnienie wyjaśniania naukowego nie jest takim zagadnieniem, w którym ten związek może się ujawniać. Jako integrację dwóch skrajnych koncepcji wyjaśniania, a mianowicie subsumpcyjną oraz przyczynową, proponuję koncepcję wyjaśniania naukowego jako subsumpcji proceduralnej. Zasadnicza jej idea sprowadza się do wyprowadzenia wyjaśniania na podstawie uogólnienia projekcyjnego.

Proponowana przeze mnie modyfikacja kryterium przyczynowości oraz proceduralna koncepcja wyjaśniania, jak pokazuję w Zakończeniu, stanowi podstawę do sformułowania alternatywnego stanowiska w stosunku do jednego z najważniejszych sporów w filozofii nauki, a mianowicie sporu realistów z antyrealistami. Jednocześnie pozwala ona prowadzić tę dyskusję niezależnie od bardzo dyskusyjnego założenia naturalizmu, które powszechnie przyjmowane jest we współczesnej refleksji nad nauką.

CZĘŚĆ I

PROCEDURALNE KRYTERIUM PRZYCZYNY

Nieeksperymentalne ustalenie przyczyny – studium przypadku

1.1 Rys historyczny odkrycia przyczyny cholery przez J. Snowa

W literaturze poświęconej odkrywaniu przyczynowości w badaniach nieeksperymentalnych wymienia się najczęściej dwa przykłady: Johna Snowa (1936) odkrycie przyczyn cholery oraz B. Hilla i R. Dolla¹ wykazanie związku między paleniem papierosów a rakiem płuc. Do tego drugiego przykładu będziemy wielokrotnie powracać w dalszej części tekstu, jednak ze względu na to, że ustalony w tych badaniach wynik opiera się także na badaniach eksperymentalnych, jako wzorcowy przykład wykrycia zależności przyczynowej w badaniach nieeksperymentalnych szczegółowiej omówię przypadek Snowa. Także i z tego powodu, że będę – podobnie jak robi to wielu innych autorów (Susser 1973, Freedman 1999) – wykorzystywał ten przykład dla opracowania probierza dla współcześnie proponowanych modeli wnioskowań przyczynowych.

Snow, jako jeszcze lekarz-stażysta, miał do czynienia z pacjentami chorymi na cholere w latach 1830. Systematycznie jednak podjął ten problem w 1849 roku, publikując pierwsze wydanie *On the mode of communication of cholera (O sposobie przenoszenia cholery)*. W momencie wybuchu epidemii cholery w Londynie w 1854 roku – gdzie pracował jako lekarz zajmujący się przede wszystkim znieczulaniem² – powrócił do tego zagadnienia

¹ Por. (Doll i Hill 1950; 1952; 1954; 1956a; 1956b; Doll i in. 2004). Kluczowe znaczenie miały też, jak podkreśla Hill, badania eksperymentalne E. Wyndera i jego współpracowników (1954; Wynder i Graham 1950; 1951; Wynder, Graham i Croninger 1953; 1955; Wynder, Bross, Cornfield i O'Donnell 1956; Wynder, Lemon i Bross 1959).

² B. W. Richardson (1936) szczegółowo opisuje rozwój kariery anestezyjologicznej Snowa, której ukoronowaniem było dwukrotne zaproszenie go do „aplikowania chloroformu” królowej podczas porodu.

z charakterystyczną dla siebie pracowitością i rzetelnością. Opracowaną wcześniej koncepcję potwierdził na szeregu nowych danych, które pieczołowicie zbierał sam lub opłacał w tym celu pomocników. Efekty tych prac zebrał w kolejnym wydaniu swojej pracy, które ukazało się w 1855 roku. W książce *Snow on cholera* (1936) znajduje się reprint tej pracy, a także późniejsze uzupełnienia autorstwa Snowa (który przeżył tę epidemię zaledwie o cztery lata).

Przed szczegółową ekspozycją tego przykładu odkrycia zależności przyczynowej należy podkreślić jego zasadniczy walor. Snow stwierdził, że przyczyną cholery jest – w zasadzie nieznanym wówczas – niewidzialny organizm żywy, który przenoszony jest głównie przez wodę. Uzasadnił istnienie tej zależności przyczynowej zarówno na poziomie teoretycznym, jak też na różnych poziomach analiz i w odniesieniu do różnych zbiorów danych i różnych populacji. Dalszego potwierdzenia tej zależności dostarczył znacznie już później R. Koch, identyfikując wskazany przez Snowa organizm jako *vibrio cholerae*.

Snow dokonał tego mimo – jak podkreśla we wprowadzeniu do *Snow on cholera* W. Frost (1936) – bardzo skromnej wiedzy i mylących danych, jakimi wówczas dysponowano: potrafiąco w zasadzie odróżniać najpowszechniej spotykane w Europie choroby zakaźne ludzi i zwierząt domowych, dość dobrze znana była ich patologia, zarówno na poziomie makro-, jak i mikroskopowym, znane były ich charakterystyczne rozkłady w przyrodzie, a także wykazano, że liczne z nich mogą być sztucznie przekazywane przez wprowadzenie do organizmu niewielkich ilości „niezdrowej materii” – to jednak nie dotyczyło infekcji układu pokarmowego, gdzie brakowało jednoznacznych dowodów empirycznych, że mogą to być choroby zakaźne.³

Natura *materies morbi* była pojmowana bardzo niejasno, choć stopniowo zbliżano się do bardziej współczesnego jej rozumienia. Powszechnie uznawano analogię między infekcją a fermentacją – tę ostatnią ponadto kojarzono z obecnością żywych organizmów o mikroskopijnych rozmiarach, ale nie wykluczano możliwości ich spontanicznej generacji. Czynniki aktywne w powstawaniu infekcji na ogół wiązano z warstwą niżej zorganizowanych mikroorganizmów, ale niekoniecznie traktowano go jako żywy organizm, a rzadziej jeszcze jako z konieczności mający naturę pasożyta „ograniczonego prawem biogenezy, jak dziś się przyjmuje” (Frost 1936, xi). Najczęściej przyjmowaną teorią wyjaśniającą związek chorób zakaźnych z mikroorgani-

³ N. Paneth (2004, 515) przypisuje Snowowi pierwszeństwo w zidentyfikowaniu oralno-fekalnego mechanizmu przekazywania i rozwoju cholery.

zmami były odmiany teorii „pythogenicznej”, upatrującej powstania tych chorób w rozkładzie materii lub nieczystości.

Cholera, jak podkreśla Frost, stanowiła szczególnie trudny przypadek do rozwiązania za pomocą:

pośredniej metody wnioskowania ze sposobu jej występowania. Trudność nie polegała na braku szczegółowych informacji. Epidemie, które przeszły przez Europę w połowie trzeciej dekady wieku, były przedmiotem ogromnego zainteresowania i były skrzętnie badane przez oficjalne komisje oraz jednostki. (Frost 1936, xi).

W Anglii, gdzie epidemie miały miejsce w latach 1831-32, 1848-49 oraz 1853-54, te studia były szczególnie drobiazgowo, zwłaszcza – za sprawą W. Farra – w odniesieniu do tych dwóch ostatnich. Szczegółowe badania i zbieranie danych prowadziły the General Board of Health – ciało powołane przez parlament, oraz the Royal College of Physicians. Jak podkreśla Frost:

Same fakty były jednak bardzo mylące. Przypadki miejscowego rozprzestrzeniania się choroby, mające pozór bezpośredniego zakażenia jednej osoby przez inną, były pomieszane z równie uderzającymi przypadkami braku zakażenia osób mających bliski kontakt z chorym oraz z przypadkami rozwoju choroby bez zauważalnych śladów związku z poprzednimi przypadkami. Ponadto, zmienność w miejscowej obecności cholery wydawała się kapryśna w najwyższym możliwym stopniu. Jeśli wydawało się, że jest regułą jej najbardziej zauważalna obecność w nisko położonych rejonach czy w okolicach składowania nieczystości, to wyjątki od takiej reguły, jak i podobnych jej, były zbyt liczne, by można je pominąć. (1936, xii)

Różne także przyjmowano teorie wyjaśniające te złożone fakty. Dominowała jednak teoria, którą w raporcie the General Board of Health, dotyczącym epidemii w latach 1848-49, opisano następująco:

Wydaje się, że jakaś materia organiczna, która konstytuuje istotę epidemii, gdy ma kontakt z inną materią organiczną pochodzącą z żywych ciał lub z rozkładu, ma moc takiego zmieniania warunków tej drugiej, że powstają w niej trujące jakości swoistego rodzaju, podobne do niej samej. (Sutherland 1854, 8; cyt. za: Frost 1936, xiii)

Odmiany tej koncepcji różniły się między sobą co do tego, czy „trucizna cholery” może być spontanicznie generowana w poszczególnych krajach, czy jest rozrzestrzenia się z kilku istniejących przed epidemią ognisk, czy jest przenoszona tylko przez rozchodzenie się w powietrzu, czy łączy się z ciałami stałymi, czy jest przenoszona przez wyziewy (łac. effluvium) chorych.

Frost jest daleki od krytyki poszczególnych teorii, jednak podkreśla, że niektóre z nich ignorowały pewne znane fakty. Wśród takich teorii znalazła się też propozycja the General Board of Health, która „stanowczo zamknęła oczy na wszystkie empiryczne dowody przekazywania [cholery] w kontaktach bezpośrednich czy na szlakach handlowych” (1936, xiii).

Po zarysowaniu okoliczności, w jakich Snow dokonał swojego odkrycia, zwróćmy za B. Richardsonem uwagę na sposób, w jaki w 1848 roku doszło do sformułowania teorii „doktora Snowa”:

Argumentował w swoim umyśle, że trucizna cholery musi być trucizną działającą na kanał pokarmowy przez wprowadzenie jej do środka i bezpośredni kontakt ze śluzową powierzchnią tego kanału, a nie przez wdychanie jakichś wyziewów. We wszystkich znanych chorobach, jak rozumował, w których najpierw zatruta jest krew, rozwijają się pewne ogólne symptomy, takie jak dreszcze, bóle głowy i przyspieszone tętno. Te symptomy poprzedzają miejscowe objawy choroby. Jednak ta reguła jest pogwałcona w przypadku cholery – symptomy są związane przede wszystkim z kanałem pokarmowym, a wszystkie późniejsze symptomy ogólnego rodzaju są skutkami wpływu z tego kanału. Wynioskował z tego, że trucizna cholery trafia bezpośrednio do kanału przez usta. Ten pogład doprowadził go do rozważenia nośnika trucizny oraz natury samej trucizny. Wiele okoliczności naprowadziło go na wodę jako na zasadnicze, choć nie jedyne, medium oraz na materię wydalaną przez pacjentów już zainfekowanych cholera, jako na tę truciznę. (Richardson 1936, xxxv)

Z publikacją tych wyników Snow wstrzymał się do chwili zebrania bardziej rzetelnych danych w 1849 roku. W dalszej części tekstu będziemy się jednak odnosić do drugiego wydania *On the mode of communication of cholera* z 1855 roku, w którym uwzględnione są obserwacje Snowa epidemii cholery w Londynie w latach 1854-1855.

Równoległe do Snowa swoje badania prowadziły (Paneth i in. 1998) The Cholera Inquiry Committee⁴ (którą w dalszej części pomijam) oraz działająca z ramienia parlamentu brytyjskiego the Board of Health. W 1848, czyli w roku drugiej fali epidemii cholery, parlament brytyjski przyjął Public

⁴ Była to komisja działająca z ramienia administracji okręgu parafii St. Jamesa (Westminster), która badała okoliczności związane z nagłym wybuchem epidemii w okolicach pompy Broad Street. Jak piszą Paneth i inni, ustalenia tej komisji potwierdziły wniosek Snowa (1998, 1545). Ponadto, ustalony został przypadek indeksowy, jakim było pięciomiesięczne dziecko mieszkające w domu przy Broad Street 40. Zmarło ono 2 września, a wodę z pranych pieluszek wylewano do wychodka, który znajdował się kilka cali od uszkodzenia w instalacji dostarczającej wodę do pompy na Broad Street (1998, 1548). Snow nie wiedział o tym przypadku, gdyż śmierć dziecka przypisano bieguncie, a nie objawom cholery.

Health Act powołujący General Board of Health na okres pięciu lat.⁵ Jednym z trzech członków tej komisji spoza parlamentu był Edwin Chadwick, autor bezkompromisowych stwierdzeń w raportach o stanie instalacji sanitarnych z lat 1848-54. Ze względu na animozje, jakie wywoływały one w środowiskach biznesowych i medycznych, parlament nie przedłużył kadencji tej komisji, która zakończyła działanie 31 lipca 1854, czyli tuż przed nasileniem się kolejnej fali epidemii.

Kolejną komisję o mniejszych jeszcze kompetencjach (wymagała corocznej autoryzacji, a w jej skład wchodziła jedna osoba będąca członkiem parlamentu) powołano w pośpiechu, już po wybuchu trzeciej fali epidemii. Benjamin Hall objął funkcję jej przewodniczącego 12 sierpnia 1854 roku, kiedy liczba zgonów spowodowanych chorobą wzrosła do 644 na tydzień. Następnego dnia odbył pierwsze spotkanie z powołanym przez siebie ciałem doradczym, jakim była Medical Council, w skład której wchodziło wielu wybitnych lekarzy, m.in. James Alderson przewodniczący the Royal College of Physicians, Benjamin Babington założyciel the London Epidemiological Society, a przede wszystkim William Farr (główny statystyk medyczny w the Registrar General) oraz John Simon (Medical Officer of Health of the City of London). Medical Council wyodrębniła trzy komisje: the Committee for Scientific Inquiries (CSI), the Committee of Treatments oraz the Committee for Foreign Correspondence. W skład CSI weszli wybitni lekarze związani z reformą sanitarną (m.in. Farr i Simon⁶). Badania naukowe dla CSI mieli prowadzić powołani przez Halla naukowcy: James Glaisher, wybitny meteorolog i założyciel the British Meteorological Society, Richard Thomson, chemik oraz Arthur Hassall, specjalista od badań mikroskopowych, znany z wykazania w latach 1840. zanieczyszczenia wody i żywności w Londynie (Paneth i in. 1998, 1545-46).

Jedną trzecią 352-stronicowego raportu CSI zajmują dane zebrane przez Glaishera (m.in. ciśnienie atmosferyczne, temperatura i jej dzienne zmiany, kierunek i siła wiatru, temperatura wody w Tamizie, opady deszczu, zachmurzenie, mgła, poziom ozonu, wilgotność, ciśnienie pary, punkt rosy – zbierane codziennie w 23 stacjach w Londynie i przedmieściach).

⁵ Do ograniczonych uprawnień tego ciała należało: kontrolowanie realizacji poprawy sieci sanitarnej zalecone przez władze lokalne, a w okresie epidemii – nadzorowanie prowadzonych akcji (Paneth i in. 1998, 1545).

⁶ J. Simon, który był w tym czasie uznawany za wiodącego eksperta w kwestiach zdrowia publicznego w Londynie, postrzegał Snowa jako obstrukcjonistę (Paneth 2004, 515).

Thomson i Hassall dokonywali niezależnych obserwacji jakości wody dostarczanej przez różne firmy oraz powietrza. Thomson odkrył obecność „vibriones” (słowo używane na określenie mikroorganizmów) tylko w próbkach powietrza z sal w całości wypełnionych przez chorych na cholera, co uznał za niewystarczającą podstawę do wyciągnięcia wniosku o związku przyczynowym:

przedwcześnie byłoby wyciągać wnioski o związku pomiędzy tą chorobą a tymi organizmami do momentu testów porównawczych na dużą skalę z innymi próbkami powietrza. (*Appendix* 1855, 131, cyt. za Paneth i in. 1998, 1547)

Hassall odkrył obecność dużej ilości „vibriones” w stolcu chorych badanym nie później niż dwie godziny po wydaleniu, jednak nie stwierdził, że są przyczyną choroby:

co daje początek czy co jest źródłem tych vibriones i jaki jest ich związek z cholera? [...] nie ma wątpliwości, że jest więcej niż jedno ich źródło. [...] Jest możliwe, że mogą dostawać się do żołądka i do jelit z atmosfery i jest doskonale pewne, że dostanie się tam często umożliwiają im niektóre z zanieczyszczonych wód, które są spożywane, a w których często odkrywałem obecność vibriones. [...] Gdy już dostaną się do układu pokarmowego, napotykają na warunki bardzo sprzyjające ich rozwojowi i przenoszeniu [...] (*Appendix* 1855, 291, cyt. za Paneth i in. 1998, 1547)

Hassall wspomina jednak, że obserwował vibriones także w stolcu zdrowych osób, chociaż w znacznie mniejszej ilości, co tłumaczył tym, że w ich przypadku nie ma warunków sprzyjających, które występują u pacjentów chorych na cholera.

Dlaczego więc CSI oraz the Board of Health odrzuciły hipotezę Snowa, przypisując większy wpływ na zachorowalność powietrzu niż wodzie? CSI dała następujący komentarz w odpowiedzi na koncepcję Snowa i zebrane przez niego informacje w odniesieniu do epidemii przy Broad Street:

Jako wyjaśnienie dużej intensywności wybuchu epidemii w bardzo ograniczonych granicach, Dr. Snow zasugerował, że prawdziwa przyczyna tego, co było specyficzne w tym przypadku, tkwiła w użyciu określonego źródła wody usytuowanego przy Broad Street w środku dzielnicy, które (jak podejrzewał) zostało zanieczyszczone odchodami chorych. Po szczegółowym badaniu nie widzimy powodu, by przyjąć takie przekonanie. Nie uznajemy za wykazane, że woda została skażona w zasugerowany sposób. Nie mamy więc wystarczających dowodów empirycznych, które wykazywałyby, że mieszkańcy tej dzielnicy, pijąc wodę z tego źródła, ucierpieli więcej niż pozostali, korzystający z innych źródeł wody. (*Report* 1855, 52, cyt. za Paneth i in. 1998, 1548-49)

Zdaniem autorów *Public health then and now* taki komentarz odzwierciedla poglądy teoretyczne CSI na mechanizm przekazywania cholery. Jak większość ówczesnych epidemiologów (z okresu poprzedzającego prace Pasteura), stanowczo sprzeciwiali się pogładowi, że cholera jest chorobą zakaźną, przenoszoną w kontaktach bezpośrednich. Cholerę przypisywali natomiast wpływowi rozkładającej się materii organicznej i oparom, które unosiły się z niej do powietrza, tzw. miazmatom. Skutek miazmatów był niespecyficzny, gdyż w zależności od okoliczności mogły prowadzić do cholery, duru brzuszego, biegunki czy tyfusu. Ogólnie więc sądzono, że cholera przenoszona jest przede wszystkim przez powietrze oraz że jej źródła nie są związane z kontaktami z innymi ludźmi. Pod wpływem J. Liebiga zarzucono prymitywną formę teorii miazmatów, zgodnie z którą stojąca woda, bagniste tereny czy rozkładająca się materia organiczna są bezpośrednią przyczyną epidemii. Uznano je bowiem za surowy materiał, na który oddziałują „fermenty” chorobowe obecne w atmosferze podczas epidemii. Te fermenty katalizowały rozkładającą się materię organiczną, wywołując śmiertelną chorobę u tych osób, które miały odpowiednie predyspozycje. Dlatego też Farr określał choroby zakaźne jako zymiotyczne od „zyme” oznaczającego ferment. Błądzący ferment nie mógł być przedmiotem bezpośredniej akcji, służącej zapobieganiu chorobie, dlatego zasadniczym sposobem było usuwanie nieczystości obecnych w środowisku. Rozumowanie kryjące się za tą koncepcją teoretyczną dobrze ilustruje następujący ustęp z wypowiedzi Farra:

Nieodkryta siła w jej [tj. cholery] błędzeniu działa w taki sam sposób, jak ferment, a więc ujawnia swoje skutki tylko w sprzyjających okolicznościach, a materiałem, z którego wywarza truciznę, musi być woda lub powietrze obfitujące w organiczne nieczystości. Biorąc to za hipotezę i testując w świetle dostępnych nam faktów, stwierdzamy, że je obejmuje i wyjaśnia. Wydaje się prawdopodobne, że infekcja może wzrastać zarówno w powietrzu, jak i w wodzie. Często trudno jest powiedzieć, które z tych mediów mogło być główną sceną trującej fermentacji, gdyż nieczystość jednej często implikuje nieczystość obu, a w większości części metropolii (gdzie najbardziej szaleje cholera) zanieczyszczenia obu rywalizują ze sobą. Uwzględniając jednak wszystkie dowody empiryczne, wydaje się niemożliwe poddawać w wątpliwość, że wpływy, jakie w całości determinują geograficzny rozkład cholery w Londynie, mniej związane są z wodą niż powietrzem. (*Report* 1855, 48, cyt. za Paneth i in. 1998, 1549)

Główny inspektor the Board of Health John Sutherland w podobny sposób wykazywał, że koncepcja zymiotyczno-miazmatyczna, jaką przyjęła oficjalnie ta komisja, może wyjaśnić mechanizm przenoszenia choroby za-

proponowany przez Snowa. Woda skażona materią organiczną jest przyczyną predysponującą⁷ cholery, w podobny sposób, jak spożycie tak skażonego pokarmu czy wdychanie tak skażonego powietrza jest przyczyną predysponującą, ale nie zawiera ona „specyficznej trucizny” („wywaru trucizny”), czyli aktywnego czynnika powodującego chorobę. Sutherland nie podejmował teoretycznej różnicy między stanowiskiem komisji a koncepcją Snowa, gdyż – jak twierdził – z punktu widzenia zapobiegania cholercie obie koncepcje prowadziły do tych samych wniosków (Paneth i in. 1998, 1549).⁸

Jak podają Paneth i in., sformułowane ostatecznie przez the Board of Health w oparciu o teorię zymiotyczno-miazmatyczną mechanizm przenoszenia cholery przebiegał następująco:

3 zbieżne czynniki: stojące powietrze z braku wiatru, wysokie ciśnienie atmosferyczne oraz wysoka temperatura wody nocą wywoływały wydobywanie się z Tamizy nocnych chmur pary zawierających nieczystości, stanowiących surowy materiał, który gdy skatalizowany przez ferment cholery obecny był w atmosferze, wywoływał nasilenie epidemii cholery. (1998, 1550)

W ten sposób też tłumaczono zaobserwowaną przez Farra zależność między przypadkami cholery a wysokością danej dzielnicy – skażone opary częściej zbliżały się do gruntu niżej położonego.

1.2 Dowody przenoszenia choroby w kontaktach bezpośrednich

Nie sposób przedstawić w całej rozciągłości bardzo szczegółowych rozważań Snowa. Stara się on bowiem wykazać istnienie, że picie zanieczyszczonej wody jest specyficzną (a nie tylko przysposabiającą)⁹ i bezpośrednią przyczyną zachorowań na cholercę. Dowody potwierdzające

⁷ Por. (Jadacki 2005, 221).

⁸ Jak podkreślają Paneth i in. nie było to zgodne z prawdą, gdyż nie stosowano wszystkich możliwych środków zapobiegających rozprzestrzenianiu się cholery, np. poprzez izolowanie chorych (Paneth 1998, 1550).

⁹ Większość ówczesnych badaczy zajmujących się cholercą dostrzegało i uznawało związek między zanieczyszczoną wodą a cholercą. Jednak zgodnie z obowiązującą teorią przenoszenia przez wdychanie niezdrowego powietrza traktowano wodę jako przyczynę, czy raczej okoliczność, która nie zawiera właściwego i aktywnego przyczynowo czynnika powodującego cholercę, lecz ułatwia jego działanie. Stąd można wyprowadzić sprzeczność między taką przysposabiającą koncepcją wpływu wody, a teorią Snowa, zgodnie z którą w wodzie obecny jest niewidzialny pasożytniczy organizm żywy, który po dostaniu się do organizmu człowieka przez pewien okres rozmnaża się przed atakiem choroby.

jego teorię pochodzą z różnych źródeł: zgodności z obowiązującymi teoriami chorób zakaźnych oraz rozmnażania się organizmów żywych, zakułowaną już wiedzą o cholery i innych chorobach, patologii cholery, mechanizmach przenoszenia tej choroby na poziomie jednostkowym oraz populacji. Ponadto, Snow wykazuje, że znane dotąd i liczne korelacje między zachorowalnością na cholere a różnymi zjawiskami czy okolicznościami (np. położeniem geograficznym, rodzajem gruntu, zawodami, ubóstwem czy zagęszczeniem mieszkańców) dają się wyjaśnić satysfakcjonująco na gruncie jego teorii, podczas gdy tylko wybrane z nich można wyjaśnić na gruncie teorii alternatywnych.

W poniższym omówieniu ograniczymy się więc do zasadniczych wątków jego rozumowania, które następnie zestawimy z innymi badaniami, jakie ówczesnie prowadzono nad epidemią cholery w Anglii, zwłaszcza komisji the Board of Health.

Swoją argumentację Snow rozpoczyna od ogólnego opisu okoliczności wskazujących, że przenoszenie cholery jest ściśle związane z przemieszczaniem się człowieka:

Podróżuje ona wzdłuż wielkich szlaków ludzkiej aktywności, nigdy nie idąc szybciej niż podróżujący ludzie, na ogół znacznie wolniej. Gdy trafia na nową wyspę lub kontynent, zawsze najpierw pojawia się w porcie morskim. Nigdy nie atakuje załóg okrętów wypływających z kraju wolnego od cholery do kraju, gdzie ona występuje, przed momentem, gdy wkroczą do portu lub gdy mają kontakt z brzegiem. (1936, 2)

Ten opis jest adresowany do tych, którzy utrzymują – zgodnie z teorią „złej materii” czy „złego powietrza” – że zasadniczym nośnikiem cholery jest „zła atmosfera”, która miałaby mieć bądź kilka źródeł poprzedzających epidemię, bądź spontanicznie powstawać w różnych miejscach epidemii.

Przechodząc do wykazania, że cholera jest przenoszona na poziomie jednostkowym w kontaktach bezpośrednich, autor *On the mode of communication of cholera* podkreśla, że te przypadki stanowią pewny dowód takiego właśnie sposobu jej przenoszenia:

Istnieje też niezliczona ilość przykładów, które dowodzą przenoszenia cholery, w odniesieniu do indywidualnych przypadków zachorowań na tę chorobę, w najbardziej przekonujący sposób. Przykłady takie, jak te poniżej, wydają się wolne od wszelkiego rodzaju błędów. (1936, 2)¹⁰

¹⁰ Jedną z interpretacji tego komentarza może być przyjęcie przez Snowa deterministycznej zależności przyczynowej na poziomie jednostkowym w następującym sensie: przy braku

Kolejny krok polega na przytoczeniu znanych z literatury przykładów, które ilustrują przekazywanie cholery w kontaktach bezpośrednich. Jako pierwszy zidentyfikowany przypadek cholery podczas epidemii w Londynie w 1848 roku – dziś tzw. przypadek indeksowy – Snow wymienia J. Harnolda, który przybył na parowcu Elbe z Hamburga, gdzie epidemia już istniała. Zmarł 22 września w kilka godzin po wystąpieniu pierwszych objawów. Kolejny przypadek miał miejsce w tym samym pokoju 30 września.

Wśród przytaczanych przez Snowa przykładów jest m.in. J. Barnes, którego nagle zachorowanie na cholere wydawało się zupełnie niewytłumaczalne do czasu, gdy jego syn wyjaśnił, że Barnes otrzymał niedawno skrzynię z niepranymi ubraniami po zmarłej na cholere jego siostrze. Wieczorem otworzył tę skrzynię, a następnego dnia wystąpiły u niego pierwsze objawy choroby (1936,4-5).

Opisy poszczególnych przypadków znane z literatury medycznej i paramedycznej Snow podsumowuje następującym komentarzem:

Łatwo byłoby, śledząc czasopisma medyczne i prace opublikowane na temat cholery, przytoczyć wiele podobnych przypadków do tych podanych wcześniej, co wypełniłoby spory tom. Jednak powyższe przypadki są zupełnie wystarczające dla wykazania, że cholera może być przekazywana od chorych do zdrowych. Jest bowiem całkowicie niemożliwe, by nawet dziesiąta część przypadków następujących po sobie kolejnych zachorowań mogła być jedynie koincydencją bez związku przyczyny i skutku. (1936, 9)

Jak podkreśla dalej Snow, zależność przyczynowa nie polega na samym kontakcie z chorym, gdyż obecność w tym samym pokoju, co pacjent, i opieka nad nim niekoniecznie wiążą się z poddaniem działaniu „złej trucizny”, a także iż można zachorować na nią, wcale nie będąc w bezpośrednim kontakcie z chorym, gdyż wywołująca chorobę „zła materia” może być przekazywana na odległość. Tym samym już te przykłady pokazują, że nie może być całkowicie słuszna powszechnie panująca teoria, zgodnie z którą cholera jest przenoszona przez wyziewy chorego, które trafiają do powietrza w pomieszczeniu, w którym przebywa, a następnie jest wdychana przez przebywających z nim.

czynników przeszkadzających wprowadzenie do układu pokarmowego „trucizny cholery” nieuchronnie prowadzi do infekcji. Inna interpretacja dotyczy determinizmu nie na płaszczyźnie zależności przyczynowej, lecz na płaszczyźnie wnioskowania. Z faktu więc, że ktoś zachorował na cholere, wynika jednoznacznie, że wcześniej wprowadził bezpośrednio do układu pokarmowego „złą materię”.

1.3 Mechanizm rozwoju choroby w organizmie jednostki

Po tych wprowadzających uwagach Snow przystępuje do szczegółowego przedstawienia alternatywnej koncepcji mechanizmu przenoszenia cholery. Podstawą dla jego określenia jest patologia tej choroby. Zdaniem Snowa, nie poprzedzają jej ogólne objawy, jak gorączka, lecz zawsze zaczyna się od objawów związanych z układem pokarmowym.¹¹ Zawsze też między momentem zakażenia a objawami upływa określony czas (zwykle od 12 do 48 godzin), co świadczy o tym, że nie jest to zatrucie substancją nieorganiczną, lecz organizmem żywym, który w tym czasie podlega intensywnej biogenezie. Obserwowane zmiany w krwi (znaczący ubytek wody w osoczu) są spowodowane przez objawy cholery (wydalanie wody), a nie – jak twierdzi Snow – przez jakiś inny mechanizm, np. obfite pocenie się (gdyż wtedy krew jest już „zmieniona”). To, że jest gęstsza i jest jej mniej, z kolei powoduje wolniejsze krążenie krwi,¹² co objawia się np. osłabieniem, omdleniem czy trudnościami w oddychaniu, a także nieprawidłowym funkcjonowaniem organów takich, jak nerki czy woreczek żółciowy. Tym samym więc Snow stara się wykazać, że wszystkie symptomy niezwiązane bezpośrednio z układem pokarmowym są efektem zmian, jakie dokonują się w krwi, a nie trucizny cholery krążącej w organizmie. Kolejnym na to dowodem ma być reakcja na wprowadzenie słabego roztworu soli do żył w momencie zapaści:

Pomarszczona skóra wypełnia się i przestaje być zimna i sina. Twarz nabiera naturalnych kolorów, a pacjent jest w stanie usiąść i przez chwilę czuje się dobrze. Jeśli te symptomy byłyby spowodowane przez truciznę krążącą w krwi i osłabiającą funkcjonowanie serca, to niemożliwe byłoby ich powstrzymanie przez wstrzyknięcie ciepłej wody z niewielką ilością węgla sodu. (1936, 13)

Snow polemizuje też z tezą, że zapaść wywołana cholera nie może być spowodowana jedynie przez biegunki i wymioty, gdyż w niektórych szczególnie trudnych i gwałtownych przypadkach ilość stolca i wymiotowanego płynu jest mniejsza niż w łagodniejszych i bardziej przewlekłych przypadkach. Snow odpowiada, że w trudnych i gwałtownych przypadkach ilość płynu, jaka została wydalona przez układ pokarmowy, jest wystarczająca do

¹¹ Nawet jeśli w niektórych przypadkach takie objawy (np. zawroty głowy, omdlenia) daje się zaobserwować, to są już skutkiem tych zmian w układzie pokarmowym, które jeszcze nie manifestują się w inny sposób (Snow 1936, 10).

¹² Jako dowody Snow przytacza też wyniki sekcji zwłok, zwłaszcza tętnic płucnych i prawej komory serca (Snow 1936, 12).

tęgo, by zaszły zmiany w krwi, charakterystyczne dla tej choroby. Natomiast większa ilość stolca występująca w innych przypadkach dowodzi tylko, że nie zatrzymana została „absorpcja z żołądka i jelit” lub że stolec był rozcieńczony wodą wypitą przez pacjenta. Ponieważ zasadnicze symptomy spowodowane są zmianami w krwi, więc to, jak poważny jest dany przypadek, zależy od tego, w jakim stopniu uszkodzona została funkcja „absorpcji” (1936, 13-14).¹³

Po scharakteryzowaniu mechanizmu choroby, Snow przechodzi do określenia tego, co jest jej nośnikiem i w jaki sposób trafia do układu pokarmowego człowieka. Ogólnie zależność tę przedstawia reguła:

Choroby przekazywane w kontaktach bezpośrednich są spowodowane przez pewien materiał, który przenosi się od chorych do zdrowych i który ma własność zwiększania się i multiplikacji w organizmach osób, które atakuje. (1936, 14-15)

W przypadku cholery nie ma „fizycznego” dowodu, jak przyznaje Snow, na to, ale pośrednim dowodem jest jej „zwiększanie się” w organizmie. A ponieważ cholera rozpoczyna się, jak wcześniej wykazywał, od ataku w układzie pokarmowym, a nie od zakażenia krwi („która nie jest pod wpływem żadnej trucizny we wczesnych etapach tej choroby”), więc „niezdrowy materiał” wywołujący cholerę musi być wprowadzony do układu pokarmowego:

musi być, w rzeczy samej, połknięty przypadkowo, bo nikt świadomie by go nie przyjmował, a wzrost ilości niezdrowego materiału czy trucizny cholery musi zachodzić we wnętrzu żołądka i jelit. (1936, 15)

Okres między momentem, gdy „niezdrowa trucizna” zostaje wprowadzona do organizmu, a pierwszymi objawami choroby określa Snow jako „okres inkubacji”, która polega na reprodukcji tego materiału we wnętrzu żołądka i jelit. Następnie

podrażnia on powierzchnię żołądka i jelit lub – co jest bardziej prawdopodobne – odprowadza płyn z krwi krążącej w naczyniach włoskowatych w stopniu analogicznym do tego, w jakim komórki nabłonka różnych organów wydobywają wydzieliny w zdrowym organizmie. (1936, 15)

¹³ Snow, powołując się na prace Parkesa i Valentina, podaje, że szacując średnią ilość krwi u zdrowego dorosłego człowieka na 30 funtów (ok. 11,2 kg), ubytek (do żołądka i jelit) związany ze zmianami wywołanymi cholerą do stanu obserwowanego przy zapaści to ok. 100 uncji (311 gr).

Dalszy wniosek Snowa dotyczy więc samego charakteru „niezdrowego materiału”, który ma wiele cech upodabniających go do organizmu żywego:

Ponieważ niezdrowy materiał cholery mający, własność reprodukowania swojego rodzaju, musi z konieczności mieć jakiś rodzaj struktury, najbardziej prawdopodobne, że takiej, jak komórka. (1936, 15)

Gwałtowność cholery w porównaniu z innymi chorobami zakaźnymi Snow tłumaczy więc tym, że jest bardzo krótki okres inkubacji (na ogół od 24 do 48 godzin) oraz ilością „niezdrowego materiału”, jaki znajduje się w tym, co wydalone jest przez pacjentów.

Snow wyraża dbałość o uzasadnienie przedstawionego mechanizmu przenoszenia cholery, podkreślając, że:

Istnieją także wystarczające dowody empiryczne, jak sądzę, na dalszych stronach tej pracy, aby udowodnić sposób przenoszenia cholery tu wyłożony, niezależnie od patologii tej choroby, ale to na podstawie rozważań dotyczących tej patologii po raz pierwszy wyjaśniono ten sposób przenoszenia [...]. (1936, 16)

Zatem zasadniczym powodem, zdaniem Snowa, rozprzestrzeniania się tej choroby jest polykanie niewielkich ilości niewidocznej, bezbarwnej i bezzapachowej trucizny cholery, znajdującej się w wydzielinach chorych. Ogromne więc znaczenie przy skali tego rozprzestrzeniania się ma brak osobistej czystości (zwłaszcza mycia rąk przed przygotowaniem posiłku i jedzeniem), a także warunki życia, gdyż, jak podkreśla Snow,

wśród klasy pracujących często posiłki spożywane są w tym samym pokoju, w którym przebywa chory, stąd tysiące przypadków w tej właśnie klasie, w których po jednym przypadku cholery w rodzinie następują kolejne, podczas gdy ludzie medycyny i inni, którzy tylko odwiedzają pacjentów, na ogół unikają zachorowania. (1936, 17)

Nawet, jak zaznacza Snow, sekcji zwłok rzadko kiedy towarzyszy zachorowanie na cholere, gdyż – w odróżnieniu od osób sprawujących opiekę nad chorymi czy mających kontakt z pościelą i łóżkiem – nigdy przy tej okazji nie są przyjmowane pokarmy i obowiązuje staranne mycie rąk.

Wśród zatem zasadniczych czynników, jakie wymienia Snow, mających wpływ na skalę rozprzestrzeniania się cholery, są: ubóstwo, związane z korzystaniem ze złej jakości wody (np. pobieranie jej bezpośrednio z Tamizy) lub jej niedostatkim, brakiem dobrego oświetlenia, a w konsekwencji niski stopień czystości, a także złe warunki życiowe, związane z niedostateczną

liczbą pomieszczeń dla wykonywania różnych czynności.¹⁴ W bogatych domach natomiast zachorowaniu na cholere przez jednego z członków rodziny często nie towarzyszą kolejne przypadki, gdyż obowiązuje zwyczaj mycia rąk oraz używa się osobnych pomieszczeń do opieki nad chorym i do jedzenia.

Po wyjaśnieniu zależności między ubóstwem a zachorowalnością na cholere, Snow przechodzi do wyjaśnienia podobnej zależności w przypadku instytucji sprawujących opiekę nad ubogimi dziećmi oraz osobami niepełnosprawnymi umysłowo. W obu przypadkach Snow podkreśla niedobór higieny osobistej, zatłoczenie oraz nieuważne wkładanie brudnych rąk do ust.

Kolejna z omawianych w *On the mode of communication of cholera* zależności dotyczy wysokiej zachorowalności wśród górników – najwyższej wśród wszystkich innych zawodów w Wielkiej Brytanii.¹⁵ Tę zależność można, jak podkreśla Snow, wyjaśnić tylko, przyjmując jego koncepcję mechanizmu rozpowszechniania się cholery. Między tym zawodem a wszystkimi pozostałymi zachodzą bowiem istotne z tego punktu widzenia różnice: w miejscu pracy nie mają żadnych urządzeń sanitarnych, muszą zabierać ze sobą jedzenie, które spożywają brudnymi rękoma. Tak więc wystąpienie cholery wśród jednego z górników w miejscu pracy jest bardzo sprzyjającym warunkiem do gwałtownego rozpowszechnienia się wśród pozostałych. Natomiast fakt, że zdarzają się przypadki górników, którzy są wywożeni na powierzchnię z zaawansowanymi objawami cholery, potwierdza Snow m.in. na podstawie własnej praktyki lekarskiej z okresu epidemii w latach 1831-32 (1936, 19-20).

Na podstawie zaproponowanego mechanizmu przekazywania cholery Snow wyprowadził też nieznanne dotąd przewidywanie, które udało mu się częściowo potwierdzić. Jeśli na cholere zachoruje osoba zatrudniona przy aprowizacji, wówczas wśród tych, którzy z niej korzystają, mogą wystąpić przypadki zachorowań. Cytując wypowiedź Johna Bloxama, opublikowaną

¹⁴ Snow wielokrotnie podkreśla związek między ubóstwem i wynikającymi z tego warunkami życia a skalą rozpowszechniania się cholery: „To właśnie wśród biednych, gdzie cała rodzina żyje, śpi, gotuje, je i pierze w jednym pokoju, cholera rozprzestrzenia się, gdy już się pojawi [...]” (Snow 1936, 18). Pozostaje pytanie, czy to wynika z troski Snowa o biednych, czy wyłącznie z pasji argumentacyjnej i chęci wyjaśnienia zależności, które były podstawą innych teorii.

¹⁵ Za N. Panethem (2004, 515) podkreślę, że oponenci koncepcji Snowa także dostrzegali związki zachorowalności na cholere z wykonywanym zawodem oraz ubóstwem, jednak: „postrzegali ubóstwo w bardziej ogólny sposób, poświęcając niewiele uwagi szczegółowym mechanizmom, za pośrednictwem których ubóstwo zwiększało ryzyko cholery”.

na łamach *Provincial Medical and Surgical Journal*, Snow zauważa, że jedyne przypadki zachorowań, jakie miały miejsce we wsi Carisbrook, związane były ze spożyciem nieświeżej wołowiny, która była własnością człowieka zmarłego wcześniej na atak cholery. Więcej szczegółów Bloxam dostarczył na życzenie Snowa:

Z tego jedzenia korzystało jedenaście osób, z czego siedem jadło je bez jakiegokolwiek dalszego przygotowania. Sześć z nich zachorowało w przeciągu dwudziestu czterech godzin po zjedzeniu, z czego pięć zmarło, a jedna wydobrzała. Siódmą osobą było dziecko, które zjadło niewielką ilość wołowiny i nie zostało zakażone. (1936, 21)

Z pozostałych czterech osób nie zaraziły się i przeżyły dwie, które jadły to mięso na ciepło, ugotowane w rosole.

Snow konkluduje, że w przypadku trudnych do wytłumaczenia źródeł zachorowań na cholere może istnieć związek z jedzeniem, gdyż często zdarza się, że sprzedawane na ulicy owoce czy inne produkty były wcześniej trzymane np. pod łóżkiem chorego na cholere.

1.4 Dowody przenoszenia choroby przez wodę

Po tych rozważaniach, dotyczących mechanizmu przenoszenia cholery związanego z bezpośrednim kontaktem z chorym, Snow przechodzi do dalszej części argumentacji:

Jeśli cholera nie miałaby innych sposobów przenoszenia niż te, które dotąd rozważaliśmy, byłaby ograniczona wyłącznie do okolic zamieszkałych przez biednych, w związku z czym byłaby skazana na przypadkowe wymarcie w jednym miejscu z braku sposobności dotarcia do nowych ofiar, ale jest często droga otwarta dla niej, aby sięgać znacznie dalej oraz by osiągnąć zasobniejsze klasy społeczności. Mam na myśli mieszanekę tego, co wydalone przez chorych na cholere, z wodą używaną do picia i celów kulinarnych, czego powodem może być bądź przesiąkanie do gruntu i docieranie do studni, bądź spływanie kanałami i ściekami do rzek, z których pobierana jest woda niekiedy dla całych miast. (1936, 22-23)

Właśnie tej części argumentacji Snowa poświęca się wspólnie najczęściej uwagi, podkreślając niekiedy, że tu wykazał się oryginalnością zasługującą na uwagę również dziś.

Omawianie poszczególnych przypadków Snow rozpoczyna od takich ilustracji, które mogą wydawać się zaskakujące, gdyż dotyczą sąsiadujących ze sobą zabudowań, które w zasadzie niczym między sobą się nie różnią poza

ilością zachorowań na cholere. Szczegółowiej omówimy tu za Snowem jeden z nich (1936, 23-24). Przypadek dotyczy dwóch dziedzińców (Surrey Buildings oraz Trusscott's Court) przy ulicy Thomas Street w Horsleydown podczas epidemii w 1849 roku, które były otoczone małymi domkami zamieszkanymi przez ubogich ludzi. W Surrey Buildings cholera okazała się szczególnie dramatyczna, a w Trusscott's Court wystąpił tylko jeden przypadek. Prowadzący badania przewodniczący Nadzoru Kanałów (the Commissioners of Sewers) stwierdził w raporcie, że jedyna różnica, jaką stwierdził między tymi dziedzińcami, była taka, że brudna woda wylewana do kanału biegnącego przed domami dostała się do studni, z której czerpali wodę, a do której woda dostarczana była przez firmę Southwark and Vauxhall Waterworks. Ponieważ pokrywa tej studni, na której stała pompa, znajdowała się na poziomie gruntu, więc okazjonalnie woda wylewała się z niej, po czym spływała z powrotem, zabierając m.in. nieczystości z kanału, który rozpoczynał się tuż przy niej. Poza tym te sporadyczne wylewy wyżyłoby niewielkie kanały w rynsztoku i chodniku, którymi stale woda spływała do tej studni, w której po opróżnieniu znaleziono znaczne składowisko nieczystości.

Pierwsze zachorowanie na cholere odnotowano w Surrey Buildings 20 lipca, kolejne następnego dnia. Ponieważ wodę do tej studni czerpano z Tamizy, więc Snow te zachorowania traktuje jako zgodne ze średnią dla tej okolicy – w sąsiednim Trusscott's Court wystąpiła podobna liczba w podobnym okresie. Jednak ponieważ w Surrey Buildings ścieki z wypranych po chorych pościeli trafiły rynsztokiem do studni, z której czerpano wodę, więc wkrótce liczba zachorowań wzrosła tam lawinowo: po tygodniu¹⁶ od tych pierwszych przypadków niemal jednocześnie u kilkunastu osób wystąpiły objawy cholery, z czego jedenaście osób zmarło.

W kolejnym z omawianych przypadków, jaki miał miejsce w Albion Terrace, Snow stara się wykazać, że nie można podać alternatywnego wyjaśnienia (1936, 25-30). W lipcu 1948 roku w siedemnastu domach odnotowano 25 zgonów, co było wówczas liczbą zupełnie nieporównywalną z sąsiednią okolicą, gdzie prawie nie spotykano jeszcze przypadków cholery. The General Board of Health podała trzy przyczyny: otwarty kanał, znajdujący się ok. 150 m od tego miejsca, z którego przy wietrze czuć było odór, po drugie odór ze zlewów w kuchniach znajdujących się na tyłach domów,

¹⁶ To opóźnienie Snow często tłumaczy okresem, który upłynął od czasu zachorowania do momentu prania pościeli po zmarłym, a następnie okresem inkubacji (1-2 dni), który poprzedza wystąpienie objawów.

a po trzecie akumulacja przypadków w jednym z domów. Snow po kolei odrzuca te czynniki jako przyczyny. Otwarty kanał oddziaływał na kilka ulic z domami, a jednak kilkanaście domów leżących między nim a Albion Terrace pozostało wolnych od cholery. Odór ze zlewów kuchennych jako powszechne wówczas zjawisko w Londynie Snow potraktował jako zupełnie nieistotne, zwłaszcza że był najsilniejszy w tych domach w Albion Terrace, w których odnotowano stosunkowo mniejszą liczbę przypadków. Z kolei akumulacja przypadków w jednym z domów nie mogła oddziaływać na inne domy na odległość. Ostatecznie więc

jest oczywiste, że jedyną specyficzną i szczegółową przyczyną związaną z ogromną plagą, jaka dotknęła mieszkańców tych domów, był stan wody, której następnie towarzyszyła cholera w prawie każdym domu, do którego była dostarczana, podczas gdy sąsiednie domy pozostały całkowicie wolne od tej choroby. (1936, 30)

Słuszność rozumowania Snowa została potwierdzona przez the General Board of Health w później opublikowanym tekście, w którym wysoką śmiertelność w Albion Terrace przypisano skażeniu wody (1936, 30).

1.5 Interwencja a ustalanie przyczyny cholery

Po omówieniu kilku kolejnych przypadków oraz relacji znanych z literatury, w których przypisywano zachorowanie na cholere wypiciu wody, Snow przechodzi do słynnej i wielokrotnie przywoływanej we współczesnych publikacjach analizy nagłego wybuchu epidemii w okolicy pompy na Broad Street w Londynie latem 1854 roku. W ciągu jednego dnia, który Snow oszacował na 31 sierpnia, zaraziło się tą chorobą ponad 300 osób.¹⁷

Snow, który mieszkał w pobliżu, pobrał pierwsze próbki wody 3 września, jednak odkryte przez niego zanieczyszczenia organiczne w wodzie były bardzo małe. Badania jej w kolejnych dniach pokazały jednak, że woda zawiera dużo takich zanieczyszczeń, tym samym uprawdopodobniając hipotezę Snowa, że do tego rodzaju zanieczyszczenia musiało dojść także feralnej nocy (1936, 39).

Snow uzyskał zgodę na zasięganie informacji o zmarłych w tej i pobliskich okolicach Londynu. Informacje, jakie zebrał o picie wody przez dotkniętych chorobą – mimo że niekompletne – potwierdzały, że jedyną cechą

¹⁷ W ciągu dwóch pierwszych dni, tj. 1 i 2 września, Snow odnotował 197 zgonów (Snow 1936, 41).

wspólną tym osobom było stałe lub okazjonalne korzystanie z pompy wodnej.¹⁸

Dla przypadków, gdzie trudno było jednoznacznie wskazać związek z wodą pochodzącą z Broad Street Pump, Snow proponował szereg mało oczywistych wyjaśnień: tę wodę powszechnie wykorzystywano w pobliskich lokalach powszechnego użytku, m.in. kawiarniach czy restauracjach, a nawet sprzedawano ją z niewielkimi domieszkami pod nazwą sorbetu (1936, 41-42).

Snow zwraca uwagę, że w kilku znajdujących się w pobliżu instytucjach średnia zgonów na cholere była zdecydowanie niższa niż w pozostałej okolicy i jedynym wyjaśnieniem, jakie podaje, jest różnica w źródle wody. Z 535 mieszkańców przytułku na Poland Street tylko 5 osób zmarło na cholere – jednak miał on własną pompę wodną i nigdy nie korzystano tu z pompy na Broad Street. Jak podkreśla Snow, „Gdyby śmiertelność w tym przytułku była równa tej na ulicach bezpośrednio go otaczających z trzech stron, to zmarłoby ponad sto osób” (1936, 42).

W browarni przy Broad Street, zatrudniającej ponad siedemdziesięciu pracowników, nie odnotowano ani jednego przypadku cholery. Korzystała ona z innych źródeł, m.in. ze studni głębinowej znajdującej się na jej terenie. Poza tym pracownicy mogli też pić sól.

Jednym z najczęściej przywoływanych przypadków indywidualnych omawianych przez Snowa jest 59-letnia wdowa, która zmarła 2 września na cholere. Mimo że nie mieszkała w pobliżu, jednak codziennie przysyłała po wodę z pompy na Broad Street. Ostatni raz piła tę wodę 31 sierpnia wieczorem i następnego dnia rano. Objawy choroby ujawniły się już wieczorem, a zmarła kolejnego dnia. Jej siostrzenica także piła u niej tę wodę i po powrocie do domu w Islington zmarła na cholere. Były to jedyne przypadki cholery odnotowane w tych okolicach (1936, 44-46).

Wielokrotnie zwraca się też uwagę na mapę, którą sporządził Snow. Dla wszystkich ustalonych miejsc zgonów podaje on ich liczbę dla poszczególnych domów znajdujących się w pobliskiej okolicy. Umieszczone zostały na niej także inne pompy, co pozwoliło Snowowi na wyciągnięcie wniosku, że

przypadki śmiertelne znacząco maleją lub w ogóle zanikają w każdym punkcie, gdzie jest zdecydowanie bliżej do innej pompy niż ta przy Broad Street. Można też zauważyć, że przypadki śmiertelne są najliczniejsze w pobliżu pompy, gdzie wodę było najłatwiej czerpać. Najbardziej ucierpiała szeroka

¹⁸ Snow osobiście zebrał dane na temat 658 domostw przy Broad Street (Paneth 2004, 515).

ulica, przy której znajduje się ta pompa, a następnie ulice od niej odchodzące, zwłaszcza te ich części, które są najbliższej pompy przy Broad Street. (1936, 47)

Kilkunaście przypadków zgonów osób z innych dzielnic wskazywał związek z Broad Street, np. jako miejscem pracy lub edukacji.

Snow sporządził także tabelę, w której umieścił liczbę odnotowanych zachorowań oraz zgonów w poszczególnych dniach. Oprócz koncentracji wokół 31 września, Snow odnotował także, że od 8 września, kiedy to na jego polecenie zdjęto ramię pompy, liczba codziennie odnotowywanych przypadków raptownie zmalała.¹⁹

Nie powiodła się próba Snowa jednoznacznego ustalenia, skąd pochodziło owo feralne zanieczyszczenie wody w pompie. Na żądanie Snowa próbki wody zostały przebadane przez wybitnego ówczesnego specjalistę od badań mikroskopowych Hassalla, który poinformował go, że zawarte w niej cząsteczki:

nie miały zorganizowanej struktury i prawdopodobnie były skutkiem rozkładu jakiejś innej materii. Odnalazł też w wodzie wielką liczbę bardzo małych owalnych drobnoustrojów, które są bez znaczenia, może jedynie jako dodatkowy dowód tego, że woda zawierała materię organiczną, dzięki której utrzymywały się przy życiu. (1936, 52)

Najbardziej wiarygodną opinię o wodzie z okresu najbliższego wybuchowi epidemii Snow uzyskał od pewnego ornitologa, który rankiem 2 września po powrocie do domu natychmiast zamówił wodę z pompy na Broad Street, którą zwykł pić. Jednak jej odrażający zapach, mimo przezroczystości, powstrzymał go od jej wypicia, co uchroniło go też przed zarażeniem, w odróżnieniu od służącej, która piła tę wodę codziennie (1936, 53).

Jako przypuszczalne wytłumaczenie, obok możliwego zanieczyszczenia z pobliskiego podziemnego kanału ściekowego lub drenu, Snow wskazuje wychodki, znajdujące się blisko i nieco powyżej studni, z której czerpano wodę do tej pompy.²⁰ Tam odnotowano kilka przypadków śmierci na cholere w okresie poprzedzającym wybuch epidemii (1936, 53).

¹⁹ Co, jak podkreśla w swoich analizach Snow, było także efektem masowej ucieczki mieszkańców tej okolicy Londynu (Snow 1936, 51).

²⁰ Snow powołuje się także na opinię „znanego inżyniera”, który twierdził, że w gliniastym gruncie, jaki znajduje się w tej okolicy, wychodki należy czyścić co 6-8 tygodni, a wiele z nich jest nieczyszczonych od dwudziestu lat, gdyż rozpuszczalna materia przesiąka do wód gruntowych (Snow 1936, 53).

Ważny, zdaniem Snowa, wniosek z tych rozważań dotyczy natury „niezdrowej materii”, która bezpośrednio powoduje cholere:

woda wszystkim wydawała się doskonale czysta i rzeczywiście zawierała mniejszą ilość zanieczyszczeń niż woda z innych pomp w tej samej okolicy, które jednak nie miały udziału w rozszerzaniu się cholery. Z tego wybuchu epidemii musimy więc wywnioskować, że ilość niezdrowej materii, która jest wystarczająca do wyprodukowania cholery, jest niepojęcie mała [...]. (1936, 53-54)

Dalszy wniosek, jaki wyciąga Snow, dotyczy alternatywnych wyjaśnień. Podany przez niego mechanizm przenoszenia cholery przez skażenie wody odchodami chorych „dostarcza dokładnego wyjaśnienia”, podczas gdy „nie ma innej okoliczności, która w ogóle dostarczałaby wyjaśnienia, niezależnie od przyjętej hipotezy o naturze i przyczynie tej choroby” (1936, 54). Takim wyjaśnieniem nie jest naruszenie pół roku wcześniej niektórych drenów, nie jest bliskość miejsca, gdzie przed dwustu laty chowano ofiary dżumy (to znajduje się w innym rejonie), ani wyziewy z kanałów, które występują we wszystkich częściach Londynu – pod względem drenów i kanałów okolice Broad Street niczym nie różnią się od tych dzielnic Londynu, które najmniej ucierpiały podczas epidemii (1936, 54-55).

1.6 Eksperyment naturalny

Przechodząc do innego poziomu analizy, gdzie dokonane zostanie zestawienie na poziomie grup mających różne źródła wody, Snow w swoim „przewidywaniu”²¹ podkreśla różnicę z dotychczas studiowanymi przypadkami:

wybuch cholery związany ze skażeniem [studni, z której czerpana jest woda do pompy lub innego ograniczonego źródła wody] jest gwałtowny i intensywny, ale także ograniczony. Natomiast gdy woda rzeki zostanie zanieczyszczona odchodami chorych na cholere, opróżnionymi ze statku lub z wpadających do niej drenów i kanałów ściekowych, przenoszenie tej choroby, chociaż na ogół mniej nagle i gwałtowne, jest znacznie bardziej rozległe, zwłaszcza gdy woda z rzeki rozprowadzana jest przez maszynę parową i rury połączone z wodociągami. (1936, 57)

²¹ To nie jest przypadek znany w literaturze metodologicznej jako „novel prediction”, kiedy nowa teoria pozwala odkryć nieznanne dotąd lub nieoczekiwane fakty. Brak bowiem analiz na poziomie grup o różnych źródłach wody traktowano jako zasadniczy brak pierwszej edycji *On the mode of communication of cholera*.

Woda z rzek jako nośnik cholery pozwala Snowowi na wyjaśnienie każdego przypadku epidemii cholery nie tylko w Londynie, ale w największych miastach Europy:²²

każda wybuch epidemii cholery w Londynie miał ścisły związek ze sposobem dostarczania wody do różnych dzielnic, modyfikowany jedynie przez biedę oraz zatłoczenie i brak czystości, które zawsze jej towarzyszą. (1936, 56-57)

W przypadku epidemii w 1832 roku Snow podaje tabelę, która ilustruje fakt, że w dzielnicy południowej Southwark, która miała najgorszą jakość wody, śmiertelność spowodowana cholera była znacznie wyższa niż w innych, zwłaszcza północnych dzielnicach, gdzie wodę dostarczano z wodociągów doprowadzonych do Tamizy powyżej odprowadzenia większości ścieków i drenów.

W dzielnicy południowej przez wiele lat rywalizowały dwie firmy: the Southwark and Vauxhall Company (połączone po 1832 w jedną firmę) oraz the Lambeth Water Company. Skutkiem tej rywalizacji było utworzenie sieci dostaw wody, która przeplatała się między mieszkańcami tej samej dzielnicy, a nawet domów na tej samej ulicy. Obie te firmy czerpały wodę z bardzo zanieczyszczonego odcinka Tamizy. Śmiertelność wśród mieszkańców korzystających z tych firm wynosiła (na 10 tys.) 120-205 w różnych dzielnicach. Chelsea, która miała podobnie ulokowany wodociąg, „poświęciła wiele wysiłku filtrowaniu wody przed jej rozprowadzaniem, dzięki czemu udało jej się oddzielić m.in. znaczną porcję tego, co powoduje cholera” (1936, 64). W przypadku tej firmy śmiertelność wahała się między 28 w Belgrave a 68 w Wandsworth. W przypadku dzielnicy Belgrave woda dostarczana przez Chelsea była jedynym źródłem, podczas gdy w Wandsworth korzystano także z pomp wodnych (1836, 66).

Snow zwraca uwagę, że zarówno jesienią 1848, jak i wiosną 1849 pierwsze odnotowane przypadki występowały w dzielnicach znajdujących się blisko Tamizy, gdzie czerpano wodę bezpośrednio z rzeki, co potwierdzałoby jego koncepcję.

Najważniejsze jednak analizy dotyczyły danych, jakich dostarczyła epidemia z lat 1854-55. W międzyczasie bowiem, tj. w 1852 roku, the Lambeth Company przeniosła wodociągi na północ, w górę rzeki, tak aby wolne były

²² Sukces, jaki Snow dostrzegł w tym wyjaśnieniu, stał się podstawą sformułowania przez niego ogólnej teorii przenoszenia chorób zakaźnych przez wodę. Taka interpretacja tylko częściowo znajduje potwierdzenie w sugestiach, jakie Snow czyni pod koniec *On the mode of communication of cholera*.

od ścieków, jakie wpływają do Tamizy w Londynie, podczas gdy Southwark i Vauxhall nadal utrzymywała swoje w okolicy Battersea Fields. Pierwsze zestawienie dotyczy różnicy w danych z roku 1853 w porównaniu z rokiem 1849 – dane, pochodzące od Registrar-General, z których korzystał Snow, nie wyodrębniają rejonu, który był obsługiwany wyłącznie przez firmę Lambeth, niezależnie od Southwark and Vauxhall. Jednak dzielnica Lambeth obsługiwana była w głównej mierze przez firmę o tej samej nazwie i śmiertelność na cholere zdecydowanie tam spadła w porównaniu z innymi dzielnicami: w tabeli z 1849 roku Lambeth zajmuje siódme miejsce (1936, 62-63), w tabeli z 1853 – trzynaste (1936, 71).

Na podstawie danych publikowanych w okresie od sierpnia 1853 do stycznia 1854 roku Snow sporządził tabelę, wydzielając poddzielnice obsługiwane wyłącznie przez firmę Southwark i Vauxhall oraz obsługiwane przez obie firmy, ale głównie przez Lambeth. W pierwszym przypadku śmiertelność na 100 tys. mieszkańców wyniosła 114, a w drugim – 60.

Ponadto, dla dwóch parafii (obiektów administracyjnych) Snow podał zestawienie dla okresu poprzedzającego przesunięcie na północ wodociągów firmy Lambeth. W 1849 roku parafia Christchurch, obsługiwana wyłącznie przez Lambeth, „miała raczej wyższą śmiertelność z powodu cholery niż sąsiadująca z nią parafia St. Saviour”, obsługiwana wyłącznie przez Southwark and Vauxhall. W 1853 roku Snow zaobserwował radykalną zmianę: parafia Christchurch miała śmiertelność 43, a St. Saviour 227 na 100 tys. mieszkańców (1936, 72).

Mimo że ta tabela dostarcza „bardzo mocnych dowodów empirycznych” istnienia „silnego związku między spożywaniem wody zawierającej ścieki miejskie a rozszerzaniem się cholery” (1936, 74), to jednak przełomowe dla obrony koncepcji Snowa okazało się uwzględnienie danych z epidemii w latach 1854-55. Klasyyczny jest już następujący ustęp z *On the mode of communication of cholera*:

pomieszczenie dostaw wody firm Southwark i Vauxhall Company z tymi Lambeth Company w znacznej części Londynu pozwoliło dokładnie przebadać temat w taki sposób, który dostarcza najbardziej niepodważalnego dowodu jednej lub drugiej stronie. W poddzielnicach, wyliczonych w powyższej tabeli jako obsługiwanych przez obie firmy, wymieszanie dostawców wody jest największe z możliwych (ang. most intimate kind). Rury każdej z firm ciągną się w dół wszystkich ulic, dochodząc do prawie wszystkich dziedzińców i uliczek. Kilka domów jest obsługiwanych przez jedną firmę, a kilka przez drugą, zgodnie z decyzją właściciela lub mieszkańca podjętą w czasie, gdy obie firmy były w okresie aktywnej rywalizacji. W wielu przypadkach pojedynczy

dom ma dostawę inną niż ten po drugiej stronie. Obie firmy obsługują zarówno bogatych, jak i biednych, domy duże i małe. Nie ma różnicy ani w warunkach życia, ani zawodach osób otrzymujących wodę z różnych firm. Jest oczywiste, że jeśli osłabienie cholery w dzielnicach częściowo otrzymujących lepszą wodę było zależne od jej gatunku, to domy otrzymujące ją powinny należeć do tych, którym w całości przypadła korzyść z osłabnięcia choroby, podczas gdy domy otrzymujące wodę z Battersea Fields nadal będą miały tę samą śmiertelność, jaką miałyby, gdyby jakoś wody w ogóle się nie poprawiła. Ponieważ nie ma różnicy ani między domami, ani ludźmi otrzymującymi wodę z tych dwóch firm, ani innymi warunkami fizycznymi, które ich otaczają, jest oczywiste, że nie można by zaprojektować żadnego eksperymentu, który mógłby bardziej gruntownie sprawdzić wpływ dostarczanej wody na postęp cholery niż to, co same okoliczności umieściły w gotowej postaci przed obserwatorem.

Ten eksperyment także był na największą skalę. Nie mniej niż trzysta tysięcy osób obu płci, każdego wieku i zawodu, każdej godności i stanowiska, od szlachty aż po bardzo biednych, zostało podzielonych na dwie grupy bez możliwości wyboru, a w większości wypadków bez ich wiedzy. Jedna grupa otrzymywała wodę zawierającą ścieki pochodzące z Londynu, a razem z nimi wszystko to, co mogło pochodzić od pacjentów chorych na cholere; druga grupa miała wodę całkowicie wolną od takich nieczystości.

Aby wykorzystać ten wielki eksperyment w wyjaśnianiu, należało jedynie poznać dostawcę wody każdego pojedynczego domu, gdzie mógł wystąpić fatalny atak cholery. (1936, 74-75)

Wziąwszy pod uwagę zarzuty, jakie anonimowy recenzent postawił pierwszemu wydaniu *On the mode of communication of cholera*, zrozumiałe jest zaangażowanie Snowa w dociekanie tej kwestii. Jednak, jak sam podkreśla,

Nie miałem powodu, by wątpić w poprawność wniosków, które wyciągnąłem z wielkiej liczby faktów już mi dostępnych, ale czułem, że okoliczność związana z wpływaniem trucizny cholery ściekami do wielkiej rzeki oraz jej rozprowadzaniem przez mile rur i nadal wywoływaniem przez nią jej specyficznych skutków była faktem tak wstrząsającej natury i o tak wielkiej wadze dla całej społeczności, że nie można było jej przebadać zbyt dokładnie czy ustalić na zbyt mocnej podstawie. (1936, 76)

Snow otrzymał dane o adresach osób umierających na cholere w odpowiednim rejonie Londynu z the General Register Office. Pozostało mu ustalenie źródła wody w każdym przypadku. Snow rozpoczął prace w połowie sierpnia. Pierwsze ustalenie do 12 sierpnia w odniesieniu do 44 przypadków śmiertelnych było zaskakujące: 38 domów otrzymywało wodę od Southwark i Vauxhall, 4 od Lambeth, a dwa korzystały z pompy. Farr, któremu Snow przekazał te wyniki, zlecił, by od 26 sierpnia obowiązkowo przekazywano

informacje o źródle wody w każdym domu w tym rejonie Londynu, gdzie odnotowano atak cholery.

Zadanie Snowa stało się znacznie prostsze, jednak, jak podkreśla, „tylko w niewielu przypadkach mogłem od razu uzyskać żadaną informację” o dostawcy wody (1936, 77). W ostateczności pobierał on próbki wody z danego domu i badał zawartość chlorku sodu (0,95 drobin w galonie wody dla Lambeth i 37,9 dla Southwark i Vauxhall).²³ Niepotrzebne nawet byłyby ścisłe obliczenia, gdyż „różnica w wyglądzie po dodaniu azotanu srebra do tych dwóch rodzajów wody była tak duża, że można było je od razu odróżnić” (1936, 78).

Snow, dzięki pomocy innego lekarza J. Whitinga, skompletował dane, z których wynikało, że w przeciągu pierwszych czterech tygodni od nasilenia się epidemii w lipcu 1854 roku (tj. do 5 sierpnia) odnotowano w odnośnym rejonie 334 zgony spowodowane przez cholere, z czego w 286 dostawcą wody była firma Southwark i Vauxhall, a w 14 Lambeth. W 22 przypadkach wodę czerpano bezpośrednio z Tamizy, w 4 z pomp, w 4 z rowów, a w 4 nie ustalono źródła (1936, 79). Uwzględniając, że pierwsza firma obsługiwała 40 046 domów, a druga 26 107, Snow otrzymał następujące wyniki przedstawione w Tabeli 1.1.

Efekt zanieczyszczonej wody, ujawniający się w około czternastokrotnie liczniejszych przypadkach zgonów, nie pozostawiał dla Snowa wątpliwości.²⁴ Zwłaszcza jeśli zestawić ten efekt ze średnią dla pozostałej części miasta. Ci, którzy otrzymywali czystsza wodę dostarczaną przez Lambeth, mieli znacznie większe szanse na uniknięcie choroby niż pozostali mieszkańcy Londynu.

²³ Snow wprowadzał do próbki wody azotan srebra, który wytrącał chlorek srebra i na tej podstawie obliczał zawartość chlorku sodu, jako wskaźnika zanieczyszczenia wody (Snow 1936, 78). Późną jesienią Snow otrzymał dodatkowe potwierdzenie słuszności wybranej metody odróżniania obu źródeł wody, gdy przekonał się, że tylko do niżej położonego jej ujęcia firmy Southwark i Vauxhall podczas przyływu dociera mieszanina wody słodkiej i morskiej (Snow 1936, 96).

²⁴ Takie wątpliwości jednak miał nawet sam Farr, który w dalszym ciągu upatrywał związku między wyżej a niżej położonymi obszarami, na których korzystano z wody. Paneth i in. (1998) pokazują jednak, że nie było do tego podstaw w danych, jakimi dysponował Farr. Inni z kolei uznawali doniosłą rolę zanieczyszczenia wody w przenoszeniu cholery, ale traktowali ją tylko jako „predysponującą” przyczynę, a nie aktywny czynnik (Frost 1936, xi).

	Liczba domów	Liczba zgonów	Liczba zgonów na 10 tys. domów
Southwark i Vauxhall	40 046	286	71
Lambeth	26 107	14	5
Pozostałe w Londynie	287 345	277	10

Tabela 1.1. (Dane na podstawie: Snow 1936, 80-81).

Ten początkowy okres był dla Snowa szczególnie ważny, gdyż zgodnie z jego koncepcją mechanizmu przenoszenia choroby największe znaczenie przy jej przenoszeniu woda miała właśnie na początku. Później bowiem, gdy liczba osób zarażonych chorobą wzrosła, istotną rolę zaczęło odgrywać przenoszenie cholery przez bezpośredni kontakt z chorymi oraz skażenie ograniczonych źródeł wody odchodami chorych. Snow pisze więc, że

Jest rzeczą niesłychanie godną uwagi, że podczas gdy tylko pięćset sześćdziesiąt trzy przypadki śmierci na cholere wystąpiły w całej metropolii w przeciągu czterech tygodni łącznie z 5 sierpnia, to ponad połowa z nich miała miejsce wśród klientów Southwark and Vauxhall Company, a duża proporcja pozostałych przypadków miała miejsce wśród marynarzy oraz osób zatrudnionych przy obsłudze statków na Tamizie, którzy prawie zawsze czerpią wodę do picia bezpośrednio z rzeki. (1936, 81)

W podsumowującym stwierdzeniu Snowa charakterystyczne jest przede wszystkim wnioskowanie kontrfaktyczne, ujawniające niezmienniczość zależności przyczynowej:

Można, w rzeczy samej, dość pewnie stwierdzić, że gdyby Southwark and Vauxhall Water Company była w stanie tak samo pospieszyć się z ukończeniem prac modernizacyjnych, jak Lambeth Company, i otrzymać wodę wolną od skażenia ściekami, to ostatni wybuch epidemii cholery byłby ograniczony w dużej mierze do osób zatrudnionych przy obsłudze statków oraz biedaków, którzy czerpią wodę pełnymi wiadrami z Tamizy lub kanałów przypiływowych. (1936, 81)

Kolejne zaprezentowane przez Snowa w tabeli dane dotyczą okresu pierwszych siedmiu tygodni (tj. do 26 sierpnia). Wszystkie dane dotyczące obszaru obsługiwanego przez Lambeth Snow zebrał, co zaznacza, sam, natomiast w pozostałej części do 5 sierpnia pomagał mu asystent, a dla pozostałego okresu oszacowano tę liczbę.

	Liczba domów	Liczba zgonów	Liczba zgonów na 10 tys. domów
Southwark i Vauxhall	40 046	1263	315
Lambeth	26 107	98	37
Pozostałe w Londynie	256 423	1422	59

Tabela 1.2. (Dane na podstawie: Snow 1936, 85-86).

Nadal więc utrzymywała się tendencja znacznie wyższej liczby zachorowań wśród klientów Southwark i Vauxhall, a wśród klientów Lambeth niżej nawet niż w pozostałych częściach Londynu obsługiwanych przez inne firmy.

W okresie od 26 sierpnia do 14 października dane pochodziły już z informacji uzyskanych przez the Registrar General (por. Tabela 1.3).

Pomijając Southwark i Vauxhall oraz Chelsea, średnia zgonów na 10 tys. mieszkańców Londynu wyniosła 34. Snow zauważa też, że mimo iż te dwie firmy mają ujęcia wody w zasadzie po przeciwnych stronach Tamizy, to jednak zastosowanie stosunkowo nowoczesnych metod filtrowania przez Chelsea (przetrzymywanie wody w zbiornikach, filtrowanie jej) dało w efekcie znacznie mniejszą zachorowalność na cholere (56 na 10 tys. osób) niż Southwark i Vauxhall (1936, 94).

Snow podsumowuje zatem, że mieszkańcy Londynu obsługiwani przez Lambeth, dostarczającą czystsza wodę niż w 1849, ucierpieli zdecydowanie mniej nie tylko niż ci obsługiwani przez Southwark i Vauxhall, ale też i średnio w Londynie oraz w pozostałych rejonach, poza centralnym i północnym.²⁵

Koncepcja mechanizmu przenoszenia cholery nie pozwalała natomiast Snowowi na prostą odpowiedź na pytanie, dlaczego epidemia nasilała się w pewnych okresach. Jej wybuch w lecie 1854 roku tłumaczył on niskimi opadami oraz wysoką temperaturą, co powodowało obniżenie poziomu wody w Tamizie i zwiększenie stężenia zanieczyszczeń.²⁶

²⁵ W kolejnej tabeli Snow przedstawia zestawienie śmiertelności na podstawie danych z 1849 roku oraz 1854 roku, wykazując zmniejszenie śmiertelności w dzielnicach obsługiwanych przez Lambeth (Snow 1936, Tabela XII, 90).

²⁶ Wyjaśnieniem tego, że dotychczas trzy wybuchy epidemii w Anglii miały miejsce latem, ma być, zdaniem Snowa, fakt, że „Anglicy z zasady nie piją wiele nieprzepracowanej wody poza okresami ciepłej pogody”, a przy posiłkach towarzyszą im inne napoje, jak herbata, kawa, likier itp. (Snow 1936, 117). Snow przyznaje, że mogą być też inne „przyczyny

	Populacja w 1851 r.	Liczba zgonów do 14 X	Liczba zgonów na 10 tys. osób
Londyn	2 362 236	10 367	43
dzielnice zachodnie	376 427	1 992	53
dzielnice północne	490 396	735	14
dzielnice centralne	393 256	612	15
dzielnice wschodnie	485 522	1 461	30
dzielnice południowe	616 635	5 567	90
Southwark i Vauxhall	266 516	4 093	153
Lambeth	173 748	461	26

Tabela 1.3. (Dane na podstawie szacunków the Registrar General podane za: Snow 1936, 88).

Kończąc analizy danych dotyczących Londynu, Snow odnosi się krótko do analiz Farra, który „odkrył zaskakującą koincydencję między śmiertelnością z powodu cholery w różnych dzielnicach Londynu w 1849 roku a poziomem gruntu” (Snow 1936, 97). Zgodnie z teorią „niezdrowego powietrza”, jaką utrzymywał Farr i większość mu współczesnych, związek między cholerą a skażoną wodą miał charakter „predysponowania”, powodowanego przez czynnik zawarty w powietrzu. Powietrze zawierające „niezdrową materię” wydobywało się np. z gruntu, stąd zaobserwowana przez niego różnica miała wskazywać, co faktycznie jest czynnikiem aktywnym w zachorowaniach na cholere. W odpowiedzi na tę koncepcję Snow zaznacza krótko, że

wzrost zachorowalności na cholere w nisko położonych dzielnicach Londynu zależy całkowicie od większego skażenia wody w tych rejonach, a w porównaniu z nimi wolna od choroby część populacji, otrzymująca lepszą wodę z Tamizy podczas zeszłego i obecnego roku epidemii, jak wykazano na poprzednich stronach, całkowicie potwierdza ten pogląd na sprawę, gdyż większa część populacji żyje w najniższych dzielnicach metropolii”. (1936, 98)

Na zakończenie swojej argumentacji Snow przytacza kilka przykładów innych miast, gdzie uwidoczniał się wpływ skażonej wody na zachorowalność na cholere, m.in. Exeter, Hull, York, Glasgow, Newcastle-upon-Tyne, Newcastle. Z zagranicznych miast natomiast: Paryż, Hamburg i Moskwę (1936, 99-109).

[...], które wspomagają rozszerzanie się cholery w okresie ciepłym” (Snow 1936, 118). Wśród takich wymienia owady, a zwłaszcza muchy.

1.7 Wyjaśnienie znanych zależności

W kończącej części argumentacji Snow odnosi się do różnic w śmiertelności w różnych grupach ludności.²⁷ Zaczyna od wyjaśnienia zależności dla płci. W pierwszym okresie po nasileniu się epidemii wyższą zachorowalność notuje się wśród mężczyzn, ale z upływem czasu liczba przypadków wśród kobiet jest wyższa.²⁸ Jako wyjaśnienie Snow podaje obowiązujące zwyczaje, mianowicie, że mężczyźni są bardziej narażeni jako ci, którzy przebywają często poza domem i jedzą w różnych miejscach. Gdy jednak epidemia się nasila, kobiety, które zajmują się chorymi oraz przygotowują jedzenie, są znacznie bardziej narażone na zachorowanie (1936, 119-120).

Z kolei przechodzi do analizy śmiertelności w poszczególnych grupach zawodowych. Najwyższa śmiertelność, jak zauważa Snow, jest wśród osób zatrudnionych przy obsłudze statków (np. marynarze, ładowacze, tragarze) oraz żyjących blisko rzeki (garbarze) i w złych warunkach (straganiarze, tkacze). Najmniej ucierpieli ci, których zawody są wykonywane w najlepszych częściach Londynu (lokaje i służba). Snow podkreśla też niską śmiertelność wśród lekarzy i grabarzy – „gdyby cholera była rozpowszechniana przez wyziewy pochodzące od pacjentów lub ciał zmarłych [...] bądź z zakażonych rejonów, to ludzie medycyny oraz grabarze byliby szczególnie narażeni na zachorowanie” (1936, 122). Także zadziwiająco niską wśród pracowników browarów, co tłumaczy niekorzystaniem przez nich z wody, dzięki czemu uniknęli zarażenia cholera tą drogą (1936, 124).

Krótki komentarz do zależności między zachorowalnością na cholera a rodzajem gruntu z pewnością nie był wystarczający do przekonania współczesnych Snowowi krytyków jego koncepcji:

Zasady, które wyłożyłem, dają zadowalające wyjaśnienie tych okoliczności, że brak drenażu wspiera rozpowszechnianie cholery oraz że rozwija się ona lepiej na gruncie gliniastym niż skalistym, piaszczystym czy żwirowym. Bez drenażu odpadki wyprodukowane przez daną populację przesiakają do gruntu i dostają się do studni, z których czerpana jest woda za pomocą pomp. [...] Skąły, piasek oraz żwir na ogół powodują puryfikację wody przez oddzielenie

²⁷ Przede wszystkim korzystając z danych Farra (Snow 1936, 119).

²⁸ (np. jesienią 1848 roku na samym początku odnotowano 80 przypadków wśród mężczyzn i 42 wśród kobiet, pod koniec jesieni odpowiednio: 258 i 210, a w pierwszym kwartale 1849 roku – 250 i 266. Latem, w tygodniu największej umieralności w lipcu, odpowiednio – 895 i 1131 (Snow 1936, 119).

lub oksydację materii organicznej, podczas gdy glina nie ma tego zbawionego wpływu w takim stopniu. (1936, 124-25)

1.8 Eliminacja hipotez alternatywnych

Kończąc, Snow podkreśla trudności alternatywnych koncepcji przenoszenia cholery. Jeśli ma być ona przenoszona przez wyziewy pacjenta do otaczającego powietrza, wdychane przez innych, to trudno jest wyjaśnić, dlaczego wiele osób do niego się zbliżających nie choruje.²⁹ Poza tym nie można wyjaśnić nagłych wybuchów epidemii, jak ten przy Broad Street.

Wyjaśnienia nie dostarcza też, zdaniem Snowa, koncepcja, która upatruje przyczyny cholery w nieznanym czynniku zawartym w atmosferze, który się lokalizuje w niektórych miejscach, a jego działanie intensyfikują gazy pochodzące z rozkładającej się materii organicznej. Takie wyjaśnienia nie mogą być zadowalające ze względu na ruch atmosfery i prawa dyfuzji gazów. Ponadto brak tu związku przyczynowo-skutkowego, gdyż w wielu miejscach Londynu silnym wyziewom nie towarzyszyła wyższa zachorowalność i na odwrót – występowała wyższa zachorowalność w takich dzielnicach, które miały otwartą przestrzeń i były czyste (1936, 113-114).³⁰

Ostatnią z omawianych teorii jest geologiczna koncepcja J. Lea, zgodnie z którą trucizna cholery istniejąca w powietrzu wokół chorego wymaga soli wapiennych lub magnezowych obecnych w wodzie pitnej, aby powstał ten efekt. Taki pogląd Snow uznaje za sprzeczny z tym, co wiadomo na temat cholery, jednak podkreśla szereg obserwacji Lea, związanych z zachorowalnością na cholere w USA, związanych z piciem wody rzecznej oraz brakiem zachorowań w przypadku picia wody źródlanej lub deszczowej (Snow 1936, 114-115).

²⁹ Snow odrzuca odwołanie się do predyspozycji niektórych z tych osób jako zadowalające wyjaśnienie na gruncie koncepcji wyziewów (Snow 1936, 114).

³⁰ Znacznie częściej niż cholera, podkreśla Snow, przykrym wyziewom towarzyszy świąd, a jednak i w tym przypadku wiadomo, że nie ma zależności przyczynowej (Snow 1936, 114).

1.9 Odpowiedź na zarzuty

Snow krótko też rozprawia się z zarzutami, które „są warte zauważenia” (1936, 111).³¹ Jeden z nich dotyczy przypadkowego wypicia wydalonej przez chorego na cholereę pacjenta wody, co pozostało bez skutku dla tej osoby. Snow odpowiedział, że „wiele warunków może być wymaganych dla przekazywania cholery, o których jeszcze nic nie wiadomo”, podobnie jak jest to w przypadku innych chorób, m.in. syfilis może być przekazany tylko we wczesnej fazie (1936, 111).

Inny zarzut, na jaki zwraca uwagę Snow, dotyczy tego, że nie wszyscy, którzy piją skażoną wodę, chorują na cholereę. W odpowiedzi podkreśla on, że w przypadku substancji organicznej, jaką jest „niezdrowa materia”, nie można oczekiwać, że będzie się ona nieskończenie rozpuszczać w wodzie, jak np. „sól chemiczna”, a zatem nie każda cząstka wody będzie zarażać (1936, 113).

Krótko także odnosi się do zarzutu, że cholera przenosi się niekiedy niezgodnie z nurtem rzek, co świadczyłoby o tym, że nie jest przenoszona przez wodę. Snow w odpowiedzi zaznacza, że wiąże się to z faktem, iż ludzie pływają także pod nurt rzeki, przenosząc cholereę na brzegach rzeki, a woda z rzeki zakaża nią osoby żyjące w danej miejscowości (1936, 124).

1.10 Profilaktyka jako test

Na zakończenie rekonstrukcji argumentacji Snowa przywołajmy dwa ostatnie tematy, jakie porusza on w swojej pracy. Jednym jest analogia między sposobem przenoszenia cholery a mechanizmem przenoszenia innych chorób zakaźnych (dżuma, febra, malaria), drugim – sposób zapobiegania cholere. We współczesnych komentarzach podkreśla się synkretyczny charakter środków zapobiegawczych, które były częściowo związane z teorią „niezdrowego powietrza”. Snow (1936, 133-35) podaje następujące zalecenia do stosowania w momencie wystąpienia epidemii: (1) zachowanie czystości przy opiece nad chorym, (2) natychmiastowe czyszczenie pościeli i ubrań chorego, (3) zapobieżenie skażeniu wody pitnej zawartością wychodków, rynsztoków czy ścieków (używanie wody z pomp tylko do mycia ulic,

³¹ Rzeczą ciekawą jest, że na tej liście nie ma zarzutu do wcześniejszego wydania *On the mode of communication of cholera*, który dotyczył zasadniczego braku dowodu przypominającego eksperyment krzyżowy.

filtrowanie wody), (4) w przypadku utrzymywania się cholery w sąsiedztwie myć lub wystawiać na wysoką temperaturę (212 st. F) wszystko, co przynosi się do domu, (5) przeniesienie chorego do osobnego pomieszczenia i pozostawienie kontaktu z nim tylko tym, którzy się nim opiekują, (6) wprowadzenie przerwy w pracy w kopalniach (co 4 godziny) na posiłek, (7) rozpowszechnienie wiedzy o mechanizmach przenoszenia choroby. Jako środki zapobiegawcze Snow wymienia (1936, 135-37): (8) drenaż, (9) zapewnienie dostaw czystej wody, (10) zapewnienie odpowiednich warunków mieszkaniowych dla biednych (gdzie są osobne pomieszczenia na gotowanie, jedzenie oraz spanie) oraz pralni publicznych, (11) kształtowanie nawyków zachowywania czystości w każdych warunkach oraz (12) ostrożność w przyjmowaniu statków przybywających z miejsc zakażonych chorobą.

Jak na zakończenie podkreśla Snow:

środki, które mają zapobiec chorobie, powinny być ufundowane na poprawnej wiedzy o jej przyczynach. Z braku tej wiedzy wysiłki, które przedsiębrano, aby przeciwstawić się cholercze, często miały przeciwny efekt. (1936, 136).

Przykładem ilustrującym wpływ nieznanomości przyczyny cholery było splukiwanie dużą ilością wody kanałów i rynsztoków, co nasiliło zachorowalność dwojako: przez wprowadzenie skażonych odchodów do systemu dystrybucji wody zanim uległy rozkładowi (i zneutralizowane zostało działanie „niezdrowej materii”) oraz przez skrócenie czasu na oczyszczenie wody przez dostarczające ją firmy, gdyż zwiększone było zapotrzebowanie na nią (1936, 136-37).

Snow wyraża nadzieję, że zachowanie odpowiednich środków zapobiegawczych, opartych na „właściwej wiedzy o sposobie przenoszenia tej choroby”, doprowadzi do „wytepienia jej z krajów cywilizowanych” (1936, 137).

2

Proceduralne kryterium przyczyny

W argumentacji Snowa na ogół zwraca się uwagę na elementy szczególnie ważne z punktu widzenia współcześnie stosowanych metod statystycznych. Tym tłumaczyłbym koncentrowanie się większości autorów przywołujących przykład badań Snowa na jego analizie nasilenia epidemii w rejonie pompy przy Broad Street, a przede wszystkim na naturalnym eksperymencie, który dotyczył zależności między jakością wody a zachorowalnością na cholere w południowych dzielnicach Londynu. Na podstawie dokonanej w poprzednim rozdziale rekonstrukcji w postępowaniu Snowa zwraca uwagę inna rzecz charakterystyczna. Mianowicie, nie ograniczał się on, jak np. Farr czy inni ówczesni badacze, do zbierania danych według przyjętego wcześniej schematu. W każdym przypadku wystąpienia cholery starał się on w sposób możliwie wyczerpujący określić wszystkie okoliczności z nim związane. Przykładem może być przywołany w poprzednim rozdziale opis zakażenia przez spożycie wołowiny. Sporządzona przez Snowa szczegółowa rekonstrukcja zdarzeń zawiera wiele elementów, które wykraczają poza wszelkie standardowo przyjęte statystyczne pomiary zjawisk towarzyszących zachorowaniu na cholere.¹

Ten charakterystyczny dla badawczego postępowania Snowa element proponuję wyodrębnić, określając go jako *interakcję badawczą* z rzeczywistością.² Sądzę, że ona właśnie ostatecznie przesądziła o tym, że Snow

¹ Szerzej zwraca uwagę na to S. Turner (1997, 36-40).

² Jak podkreślam w Zakończeniu, jest to nawiązanie do tradycji szkoły lwowsko-warszawskiej; por. (Kamiński 1976/1994, 421 p. 4). Także do pewnych intuicji języka polskiego: „[...] zauważyć warto, że w pojęciu wywoływania czy stosunku przyczynowego tkwi jakiś element aktywności, stąd chętnie nazywamy przyczynę ‘czynnikiem’. Przyczyna działa,

ostatecznie odkrył przyczyny cholery i je uzasadnił. Podkreślmy, że analizując uzyskane w tej interakcji dane o zjawisku cholery i zjawiskach mu towarzyszących, odnosił się w pierwszym rzędzie do danych pochodzących z bezpośredniej interakcji z rzeczywistością, starannie oddzielając je od danych niekompletnych lub tylko szacunkowych.

Postępowanie badawcze Snowa wskazuje więc drogę pośrednią między dwiema skrajnościami, które krótko wysłowił J. Gutenbaum (1987, 86):

Jednakże doświadczalne potwierdzenie związku przyczynowego nie może być dokonane poprzez obserwację bierną, lecz jedynie drogą eksperymentu czynnego.

Bierna obserwacja, jak podkreśla Gutenbaum (1987, 88), może jedynie pełnić funkcję negatywną i służyć obaleniu związku przyczynowego.

W dalszej części tekstu pojęcie interakcji badawczej traktuję jako pojęcie pierwotne na gruncie metodologii i filozofii nauki. W oparciu o pewne takie jej własności, jak: zorientowanie na daną dziedzinę rzeczywistości, ukierunkowanie teoretyczne oraz niezmienniczość, określam w kolejnych dwóch punktach warunki, które muszą spełniać propozycje kryteriów przyczynowości, czyli tzw. *warunki adekwatności* dla tych kryteriów. W Punkcie 2.3 natomiast formułuję propozycję pewnego kryterium przyczynowości, które te warunki adekwatności zachowuje, a które proponuję określać jako kryterium proceduralne.

2.1 Warunki adekwatności kryterium przyczyny

W przypadku badania naukowego interakcja przyczynowa jest ściśle *zorientowana* na daną dziedzinę rzeczywistości D, w której wyodrębnia się część zjawisk jako przedmiot badania.³ Często te zjawiska reprezentowane

jej skutek jest rezultatem tego działania” (Gumański 2004, 114). Interesujące byłoby też odniesienie pojęcia interakcji badawczej do „praktyki naukowej” Krajewskiego, której „główną metodą” jest eksperyment, ale które obejmuje także obserwację, wymagającą „wielu dodatkowych czynności, tworzących korzystne warunki obserwacji. Jest więc pewnym oddziaływaniem na rzeczywistość” (1998, 63-64). Por. też (Wójcicki 1995, 33).

³ Parafrazując pytanie P. Feyerabenda, można by zapytać, czy możliwa jest interakcja bez doświadczenia (1969/1979, 194-99). Prowadziłyby to jednak – podobnie jak dzieje się to w dyskusji Feyerabenda – do rozważania możliwości, które nawet po osiągnięciu kresu nauki byłyby hipotetyczne.

są w postaci zbioru zmiennych V ,⁴ a zasadnicze zadanie, jakie stawia się interakcji przyczynowej z tą dziedzina polega na ustaleniu zależności L zachodzących między – wszystkimi lub wybranymi – badanymi zjawiskami.

W przypadku badań Snowa przedmiotem interakcji była epidemia cholery badana w aspekcie epidemiologicznym, a więc mechanizm przenoszenia tej choroby. Z tej perspektywy uwzględnił on szereg zjawisk, które dotychczas wiązano z tą chorobą: zanieczyszczenia atmosferyczne (głównie lokalne), wyziewy chorego, pory roku, klimat, wysokość i rodzaj gruntu, płeć, zawody, status społeczno-ekonomiczny, poziom higieny, ilość mieszkańców przypadających na jeden dom, warunki życia (przebywanie chorego w jednym pomieszczeniu z pozostałą częścią rodziny) czy istnienie i jakość sieci wodno-kanalizacyjnej. Uwzględnił także zjawiska znane, dotyczące patologii cholery: brak ogólnych objawów poprzedzających specyficzne dla cholery, biegunka, skurcze żołądka i jelit, wymioty, gorączka, odwodnienie, osłabienie czy zapaść. Zwrócił jednak uwagę na nowe zjawiska: w odniesieniu do patologii to przede wszystkim zmiany w krwi, w odniesieniu do mechanizmu przenoszenia choroby to przede wszystkim jakość dostarczanej wody, skażenie pokarmów lub napojów (odchodami lub innymi substancjami wydzielonymi przez chorych) oraz – dość ogólnikowo – na zwyczaje kulinarne (Anglików i Szkotów).

Podkreśla się (Susser 1973; Freedman 1999) także, iż Snow na różnych poziomach analizy (dla różnych rodzajów jednostek): od indywidualnego poprzez poziom gospodarstw domowych po grupy społeczne wykazał spójność związku przyczynowego dla jednostek tych różnych poziomów.

M. Susser w klasycznym podręczniku z zakresu epidemiologii *Causal thinking in the health sciences* (1973, 54-56) zwraca uwagę szczególnie na przejście od danych, jakie Snow otrzymywał od Farra, ówczesnego the Registrar General, na poziomie grup społecznych (m.in. liczba zgonów w poszczególnych dzielnicach Londynu, położenie nad poziomem morza, za-
możność, gęstość zaludnienia oraz dostawca wody) do danych na poziomie poszczególnych gospodarstw domowych (zwłaszcza o źródle wody, z którego korzystała osoba chora lub zmarła na cholere).⁵ Przypomnę, że kiedy latem 1855 roku nasiliła się epidemia cholery w Londynie, Snow miał już opracowaną teorię jej rozwoju w organizmie człowieka i mechanizmu prze-

⁴ Najczęściej spotykane w badaniach naukowych rodzaje zmiennych charakteryzuje J. Brzeziński (2000, 28-55).

⁵ Na złożoność wzajemnej zależności badań Snowa i Farra zwraca uwagę Morabia (2001).

noszenia tej choroby. Gdy uświadomił sobie, jak dalece system dostaw wody z różnych ujęć jest wymieszany w południowych dzielnicach Londynu, poprzez podjęcie interakcji z tym zjawiskiem, postanowił uzyskać dokładne informacje o dostawcy wody dla każdego zarejestrowanego przypadku wystąpienia cholery w tych poddzielnicach Londynu. Po otrzymaniu zgody na pozyskiwanie adresów osób zmarłych, pieczołowicie ustalał źródło wody, z którego korzystała osoba mieszkająca pod danym adresem. Ustalone przez Snowa zależności w ciągu czterech pierwszych tygodni badań skłoniły Farra do wprowadzenia obowiązku od 26 sierpnia 1855 roku sprawozdawania dla każdego przypadku osoby zmarłej, z jakiego źródła wody korzystała.

Jak pokażę szerzej na przykładzie równoległych do Snowa badań nad epidemią cholery w Londynie, nie każda interakcja zorientowana na daną rzeczywistość prowadzi do ustalenia zależności przyczynowych między zmiennymi. Pozostaje więc zasadnicze pytanie, kiedy interakcja przyczynowa prowadzi do ustalenia zależności przyczynowych między badanymi zjawiskami? Nie ma, jak się wydaje, jednoznacznej odpowiedzi na to pytanie, z czym wiąże się także brak formalizacji tych kryteriów. Jak bowiem podkreśla D. R. Cox:

istnieje szeroka gama możliwości pomiędzy ogólną wiarygodnością (ang. plausibility) na gruncie wiedzy o danym przedmiocie badań a pełnym zrozumieniem opartym na szeroko akceptowanych zasadach i bezpośrednich dowodach empirycznych. (Cox 1996, 225)

Takie kryteria zresztą są formułowane dla zaproponowanych modeli, a nie – jak tutaj się proponuje – dla bardziej podstawowej interakcji przyczynowej ze światem. Zasadnicze z takich kryteriów dla zorientowanej interakcji w badaniach naukowych jest jej *ukierunkowanie teoretyczne*.⁶ Jeżeli interakcja zorientowana jest na dziedzinę D, w której wyodrębnione są zjawiska V, to ukierunkowanie teoretyczne tej interakcji można rozumieć jako wykluczenie możliwych zależności między zjawiskami V. Takie wykluczenie może być oparte na wiedzy teoretycznej lub – w przypadku gdy ta nie

⁶ Por. (Giedymin 1964, 12-16; Nowak 1985, 46-58; Wójcicki 1995, 38; Brzeziński 1996, 37-44). Problematykę ukierunkowania teoretycznego interakcji badawczej – a zwłaszcza problemu niezdeteminowania wszystkich relewantnych zjawisk oraz wyboru zjawisk opisywanych w teorii – szczegółowo ilustruje na przykładzie powstania modelu atomu N. Bohra W. Sady (2003, 112-18).

jest dostępna – na posiadanej ogólnej wiedzy o danej dziedzinie lub dostępnych dowodach empirycznych jej dotyczących.⁷

Wprowadzenie takiego wykluczenia może dokonać się co najmniej dwojako:

- (1) Po ustaleniu relacji między badaną sferą zjawisk V w danej dziedzinie rzeczywistości a inną sferą zjawisk V' dla tej samej dziedziny i po podaniu teoretycznych lub empirycznych zależności L' dla V' wykluczenie spośród zależności L dla zjawisk V te, które są niezgodne z L' (kontrakcja).
- (2) Uwzględnienie dotychczas zaproponowanych ograniczeń zależności L między zjawiskami V i wykorzystanie wiedzy teoretycznej o danej dziedzinie lub dowodów empirycznych (dla zjawisk V lub V') wykluczenie tych ograniczeń, które nie są zgodne z tą wiedzą lub dowodami empirycznymi (ekspansja).

Przykładem kontrakcji może być uwzględnianie dziś nie wyartykułowanej przez Snowa zależności między zachorowalnością na cholera a odpornością organizmu czy pewnymi własnościami układu trawiennego. Ilustracją ekspansji natomiast może być wyeliminowanie dostrzeżonej przez Snowa – tylko ogólnikowo – zależności między sezonowością zachorowań na cholera a obyczajami kulinarnymi różnych nacji.

Te dwa sposoby wykluczania możliwych zależności między zmiennymi nie są wykluczające, lecz komplementarne, gdyż możliwe jest podanie takich teoretycznie lub empirycznie ugruntowanych zależności L' dla zjawisk V' , które jednocześnie pozwolą na wykluczenie szeregu zależności dla zjawisk V oraz wykluczenie innych teoretycznych ograniczeń tych zależności, które są niezgodne z L' .

Zapytajmy więc, w czym ukierunkowanie teoretyczne interakcji różni się od tego, co postuluje J. Goldthorpe, a mianowicie od hipotetycznego przyjęcia procesu przyczynowego? Przypomnę krótko, że charakteryzując postępowanie badawcze w naukach społecznych zmierzające do wyodrębnienia zależności przyczynowych, Goldthorpe wyróżnił trzy etapy (2000, 151): (i) ustalenie zjawisk, które stanowią eksplananda, (ii) wprowadzenie hipotezy o procesie generującym obserwowane zjawiska, który funkcjonuje na poziomie działań społecznych jednostek oraz (iii) testowanie tych hipotez.

⁷ R. Braithwaite formułuje podobną tezę, wychodząc od krytyki tradycyjnej empirystycznej koncepcji przyczynowości i twierdząc, że dla odróżnienia związku przyczynowego między X a Y od zwykłej koniunkcji należy przywołać hipotezę ogólną, na podstawie której eksplikuje się tę różnicę (Braithwaite 1955, 317).

Może się zatem wydawać, że zaproponowane tu ukierunkowanie teoretyczne jest w zasadzie tożsame z wprowadzeniem hipotezy o procesie generującym zależności L między zjawiskami V. Podkreślony przez Goldthorpe'a warunek, by te hipotezy dotyczyły poziomu działań społecznych oznaczałby dodatkowo, że postulując dla badanej dziedziny proces generujący zjawiska V, należy odwołać się do zjawisk V' na „poziomie hierarchii niżej” (Cox, Wermuth, 2001, 69).

Goldthorpe przyznaje jednak (2000, 156), że obserwacje nie są z poziomu niższego, np. działań poszczególnych jednostek, lecz wyższego – grup społecznych. Żeby więc taki proces generujący miał odniesienie do zaobserwowanych zjawisk, autor *On sociology* przywołuje inny aspekt analiz, mianowicie aspekt diachroniczny: jeżeli taki proces rzeczywiście generuje dane, to możliwe jest

zaplanowanie takich mniej bezpośrednich testów przez określenie innych efektów, które powinien on wygenerować oprócz tych, które składają się na wyjaśniane przez niego regularności, a które też można stwierdzić empirycznie. (Goldthorpe 2000, 156)

Skutki takiego więc procesu generującego nie są oceniane w kontekście (pojedynczego) randomizowanego eksperymentu, lecz w kontekście

teoretycznie ukierunkowanego wyjaśnienia (ang. informed account) procesu generującego, który jest przedmiotem ciągłej oceny, a jej wynikiem jest falsyfikacja lub, jeśli test jest pozytywny (ang. withstood), po prostu korroboracja raczej niż weryfikacja skutków w rodzaju „raz-na-zawsze”. (Goldthorpe 2000, 156)⁸

Jak subtelnie podkreśla Cox, takie hipotetyczne wprowadzenie procesu generującego w kontekście synchronicznym

By było zadowalające, musi być związane z istnieniem dowodów empirycznych, zwykle uzyskiwanych w różnego rodzaju badaniach, że takie generujące procesy nie są tylko hipotetyczne. Przyczynowości nie można ustalić przez tylko nazwanie statystycznego modelu przyczynowym. (Cox 2001, 69)

W odniesieniu więc do zaproponowanego przez Goldthorpe'a hipotetycznego wprowadzania procesów generujących, można powtórzyć zarzut empirysty: w propozycji Goldthorpe'a przyczynowość nie podlega ocenie zgod-

⁸ Jak zostanie to wyjaśnione pełniej poniżej, należy uznać wnioski Goldthorpe'a, że – przynajmniej nie zawsze – eksperyment z randomizacją stanowi optymalny kontekst dla oceny i uzasadnienia zależności przyczynowej między zmiennymi. Jednak powody, które podaje metodologia proceduralna, dla tego wniosku są zasadniczo odmienne od Goldthorpe'a.

ności empirycznej z zaobserwowanymi zjawiskami. Wprowadzając aspekt diachroniczny daje natomiast podstawy pragmatycznej *akceptacji* tego procesu przyczynowego, czyli gotowości posługiwania się nim z *innych powodów niż prawdziwość* określana w odniesieniu do dostępnych danych.

Kolejny zasadniczy brak w postulowanym przez Goldthorpe'a hipotetycznym wprowadzaniu procesów generujących dotyczy tego, że nie określa on nie tylko dokładnych kryteriów przyczynowości, ale w ogóle nie pokazuje, jak w aspekcie diachronicznym akceptacja hipotetycznie wprowadzonego procesu przyczynowego wiązałaby się z prawdziwością.⁹ Co ostatecznie w aspekcie diachronicznym, rozstrzygałoby o prawdziwości jednej hipotezy względem innej, nawet jeśli dopuścić nieskończoną liczbę powtórzeń zbierania danych?

Zgodnie z przytoczonym zaleceniem Coxa, ukierunkowanie teoretyczne interakcji nie może w żadnym wypadku zaniedbać kwestii zgodności z wiedzą uzyskaną w dotychczasowych interakcjach z doświadczeniem, odwołując się do możliwej przyszłej zgodności. Innymi słowy, nie można – jak wydaje się postulować Goldthorpe – zaniedbać zgodności z doświadczeniem w aspekcie synchronicznym (tu i teraz), twierdząc, że taka zgodność dla procesu generującego uzyskana będzie w aspekcie diachronicznym (kiedyś w przyszłości).¹⁰ Wyjaśnienie tej okoliczności jest następujące. Załóżmy, że dzięki dotychczasowym interakcjom posiadamy pewną wiedzę o zależnościach między zjawiskami $O \subseteq V$ danej dziedziny D , a także pewną wiedzę teoretyczną o tych zjawiskach.¹¹ Jeżeli przyjęty zgodnie z postulowanym przez Goldthorpe'a sposobem hipotetyczny proces przyczynowy jest jedynym zgodnym z tą wiedzą teoretyczną, wówczas w zasadzie w jego akcepta-

⁹ Takim związkiem mogłoby być np. wykazanie zbieżności modelu opartego na postulowanym procesie przyczynowym do prawdziwego modelu przy np. danych (rozmiarach próbki) rosnących nieskończenie. Nie widać jednak, dlaczego – gdyby nawet istniał taki dowód – miałoby się to w jakiś zasadniczy sposób łączyć z własnościami postulowanego procesu, a nie z samym modelem zależności między zmiennymi.

¹⁰ W podanym przez Goldthorpe'a opisie testowania hipotezy o procesie przyczynowym (Goldthorpe 2000, 157) nie zaznacza się aspekt diachroniczny: zamiast wprowadzać do analizy współzmiennie wpływające na dzieci, które utraciły i nie utraciły rodziców wskutek rozwodu, czego wymagałoby zdaniem Goldthorpe'a manipulacyjna koncepcja przyczynowości jako wyidealizowanego eksperymentu, testowanie procesu wymagałoby uwzględniania jego konsekwencji dla przypadków utraty rodziców z innych powodów itp. (Goldthorpe 2000, 157).

¹¹ Jako oczywisty przypadek pomijam tu sytuację akceptacji hipotetycznego procesu przyczynowego, który byłby niezgodny z wiedzą teoretyczną lub wiedzą o zależnościach między zjawiskami D .

cji aspekt diachroniczny, na który zwraca uwagę Goldthorpe, nie ma specjalnego znaczenia. W tym przypadku więc ukierunkowanie teoretyczne interakcji będzie zbieżne z tym, co Goldthorpe określa jako postulowanie hipotetycznego procesu generującego.

Jednak w przypadku, gdy postulowanie hipotetycznego procesu generującego wiąże się ze zignorowaniem innych możliwości zgodnych z wiedzą o zależnościach między badanymi zjawiskami ze względu na oczekiwaną zgodność w aspekcie diachronicznym, koncepcja Goldthorpe'a wydaje się problematyczna. W tym przypadku bowiem w postulowanym tu sposobie akceptacji procesu przyczynowego ustalenie zależności przyczynowych dla badanych zjawisk dokonywałoby się nie poprzez interakcję z badanym światem ukierunkowaną teoretycznie, lecz niezależnie od niej. Ponadto, wzięwszy pod uwagę fakt, że dla pewnych dziedzin można *ex post* zawsze podać wyczerpujące wyjaśnienie (Cox, Wermuth 1996, 226)¹² oraz że zawsze ocena zgodności z danymi dokonywana jest synchronicznie i w oparciu o dostępny zbiór danych (Cox, Wermuth 2001, 69), należy uznać, że postulowany przez Goldthorpe'a sposób akceptacji generujących procesów przyczynowych może prowadzić do akceptacji jako procesów przyczynowych takich hipotez, które w żadnym sensie nie są przyczynowo związane z badaną rzeczywistością.

Wbrew pozorom zbieżności proponowanej tu koncepcji ukierunkowanej teoretycznie interakcji z postulowaną przez Goldthorpe'a akceptacją hipotetycznych procesów przyczynowych zachodzi między nimi zasadnicza różnica. Jak starałem się wykazać, zalecenia Goldthorpe'a okazują się bowiem bądź mało specyficzne, bądź zasadniczo rozbieżne w stosunku do koncepcji ukierunkowanej interakcji.

Ukierunkowanie teoretyczne interakcji przyczynowej z badaną dziedziną rzeczywistości jest warunkiem koniecznym, ale niewystarczającym odkrycia zależności przyczynowych między badanymi zjawiskami. Zilustrujmy to na przykładzie badań nad epidemią cholery w Londynie w XIX wieku. Snow, jako odbywający praktykę lekarz, zetknął się po raz pierwszy z pacjentami chorymi na cholere podczas epidemii 1831-32 w Newcastle-upon-Tyne. Jak wynika z jego relacji, potwierdzonej przez Richardsona, kwestię przyczyn cholery podjął dopiero w 1848 roku, w związku z nową falą epidemii. Jako lekarz, mający wcześniejsze doświadczenie w leczeniu chorych na cholere, podjął interakcję ze zjawiskiem epidemii na poziomie poszczególnych przy-

¹² A raz zaakceptowany hipotetyczny proces przyczynowy odpowiednio do danych modyfikować.

padków, stosując obie wskazane wcześniej strategie przy teoretycznym ukierunkowywaniu tej interakcji. Snow wykazał, że na poziomie patologii cholery, obserwowanej w poszczególnych przypadkach zachorowań, zasadnicza alternatywa, jaką była teoria wyziewów, nie dostarcza spójnego wyjaśnienia tego, co znane było z dotychczasowych interakcji oraz badań zmian krwi podjętych przez Snowa. Przypomnę krótko, że zgodnie z wysuniętą koncepcją przez Snowa, zakażenie związane jest z wprowadzeniem bezpośrednio do układu pokarmowego mikroorganizmu, który następnie po okresie wzmoczonego wzrostu (inkubacji) atakuje organizm chorego, a następnie wydostaje się wraz z odchodami i kontynuuje proces zakażenia innych. Konsekwencją takiej koncepcji rozwoju cholery w organizmie człowieka jest to, że nie ma ogólnych symptomów choroby, które poprzedzałyby symptomy charakterystyczne dla cholery (i związane z układem pokarmowym), oraz że zakażenie nie dokonuje się za pośrednictwem krwi, w związku z czym zmiany w krwi są obserwowalne w końcowym okresie inkubacji. Tym samym teoria wyziewów, która przewidywała zakażenie poprzez powietrze, a więc poprzez krew, dawała w tym względzie dwa przeciwstawne przewidywania i nie wyjaśniała ustaleń dokonanych w interakcji ze zjawiskiem cholery.

Taki mechanizm rozwoju, jaki związany był z koncepcją Snowa, prowadził do wyodrębnienia dwóch zasadniczych mechanizmów przenoszenia cholery: bezpośrednie wprowadzenie mikroorganizmu ze skażonym pokarmem oraz ze skażoną wodą. Działanie obu tych mechanizmów było modyfikowane przez poziom higieny oraz związany z warunkami życia i statusem społeczno-ekonomicznym kontakt chorego i jego rzeczy z żywnością i innymi domownikami.¹³ Skażenie poprzez bezpośrednie przeniesienie odchodów chorego na spożywaną żywność pozwalało Snowowi na wyjaśnienie lokalnych przypadków zachorowań na poziomie poszczególnych jednostek (zwłaszcza przenoszenie w kontaktach bezpośrednich) oraz – także lokalnie – na wyższych poziomach. Jak podkreślał Snow, takie oddziaływanie, które wymaga bezpośredniego kontaktu z odchodami chorego, jest zawsze ograniczone w przestrzeni i czasie.

Drugi więc mechanizm, jakim było wprowadzenie mikroorganizmu wraz ze skażoną wodą, pozwala nie tylko na wyjaśnienia lokalne, ale także na objęcie nim szerszej skali przestrzeni i czasu.

¹³ Por. też (Adams i in. 2003a, 2003b; Adda i in. 2003; Poterba 2003).

Oba mechanizmy przenoszenia są spójne z mechanizmem rozwoju choroby i dlatego Snow mógł, zachowując spójność, posługiwać się obydwoma w wyjaśnianiu wszystkich obserwowanych ówczesnie zależności dla epidemii cholery. Nie wszystkie z tych wyjaśnień przetrwały próbę czasu, czego chyba najbardziej jaskrawym przejawem jest wyjaśnienie zależności między nasileniem epidemii a porami roku. Snow odwołuje się tu do nawyków Anglików, którzy jesienią piją mało wody, a latem dużo. Koncepcja Snowa nie przewidywała bowiem możliwości skażenia wody w inny sposób niż w koniecznym związku z odchodami chorych.¹⁴

Szczególnie wiele uwagi Snow poświęca wyjaśnieniom – na wielu diskutowanych poziomach analizy zjawisk – przypadków, gdy skażona woda była jedynym czynnikiem różniącym między przypadkami zachorowań a brakiem zachorowań (np. pompa przy Broad Street, dostawcy wody dla południowych dzielnic Londynu).

Oba wymienione wcześniej rodzaje wykluczeń zależności między badanymi zjawiskami, a więc oparte na niezgodności między różnymi poziomami analiz oraz niezgodności z posiadaną wiedzą teoretyczną lub empiryczną, wykorzystuje Snow w argumentacji przeciw alternatywnym teoriom.

Najbardziej szczegółowa alternatywa, jaką zaproponował Farr na podstawie danych z epidemii w latach 1848-49, przewidywała związek między wysokością położenia danej dzielnicy nad poziomem morza a ilością przypadków śmiertelnych (zdecydowana większość przypadków była odnotowana w dzielnicach południowych). Taka zależność jednak, jak pokazywał Snow, nie tłumaczyła wielu innych znanych zależności, między innymi zróżnicowania pod względem zachorowalności na cholere w dzielnicach północnych, które różniły się tylko źródłem wody.

Powracając zatem do dyskusji kryteriów interakcji, która wiedzie do ustalenia zależności przyczynowych między zmiennymi, można z tego przykładu wyprowadzić następujące wnioski. Członkowie komisji w różny sposób podejmowali interakcję z różnymi zjawiskami w zakresie cholery. Powstaje jednak pytanie, czy ta interakcja była zorientowana na badaną dziedzinę rzeczywistości?

Wyodrębniony został cały szereg zmiennych, które badano, w tym zmiennych związanych z pogodą, zanieczyszczeniem wody pochodzącej z różnych źródeł oraz powietrza w różnych okolicach Londynu. Jak podkre-

¹⁴ Współcześnie nasilenie cholery wraz z ociepleniem tłumaczy się m.in. nasileniem rozwoju mikroorganizmów żyjących w wodzie, na których pasożytuje *vibrio cholerae* (Colwell 1996, 2030).

ślają jednak Paneth i jego współpracownicy, komisja w żaden sposób nie odnosiła zymiotyczno-miazmatycznej teorii do zjawisk znanych z patologii cholery, a zwłaszcza podkreślanego przez Snowa związku między utratą płynów z organizmu a zapaścią naczyń krwionośnych.

Brak odniesienia teorii zymiotyczno-miazmatycznej do rozwoju cholery w organizmie pacjenta należy więc scharakteryzować nie jako problem braku zorientowania interakcji, lecz niespełnienia kryterium, na które zwrócę uwagę w kolejnym punkcie. Nie można bowiem zarzucić badaniom prowadzonym przez komisję braku określonego związku z badaną dziedziną rzeczywistości, jaką była epidemia cholery. Mimo że sporą część zjawisk, które uwzględniano, dziś można by uznać za niezwiązane z tą dziedziną rzeczywistości, to jednak taki związek był wprowadzony poprzez teorię zymiotyczno-miazmatyczną, którą posługiwała się komisja.

Aż nadto oczywiste wydaje się także ukierunkowanie teoretyczne interakcji podejmowanych przez członków komisji. Przejawem tego jest nie tylko uwzględnienie bogatej palety zjawisk, ale także wyjaśnienie szeregu znanych zależności w oparciu o zaproponowany na gruncie tej teorii mechanizm. Ten mechanizm tłumaczył bowiem nagłe nasilenie się epidemii wraz z nadejściem ciepłej pory roku, czego w zadowalający sposób nie potrafił wyjaśnić Snow.

Mimo tego, że interakcja podjęta przez komisję spełnia warunki zorientowania i ukierunkowania teoretycznego, to jednak koncepcja, jaką zaproponowała, nie przetrwała próby czasu. Można więc wyprowadzić z tego wniosek, że te dwa wcześniej zaproponowane warunki nie są wystarczające do uzyskania zależności *przyczynowej* między zmiennymi.

2.2 Niezmienniczość przyczyny

Proponuję więc poniżej zająć się kolejnym warunkiem adekwatności. Jak stwierdziłem powyżej, komisja – w przeciwieństwie do Snowa – nie podjęła próby wyjaśnienia mechanizmu rozwoju choroby na poziomie organizmu jednostki. Taka próba na gruncie koncepcji Snowa nie tylko wykazywała jej spójność, ale przede wszystkim wskazywała na inną cechę, którą chcę tutaj wyartykułować. Zależność przyczynowa między trucizną cholery wprowadzoną do organizmu np. przez skażoną wodę a zachorowalnością okazała się w rozważaniach Snowa *niezmiennicza* w tym sensie, że można było stwierdzić jej zachodzenie mimo zmiany poziomu analizy zjawisk poprzez wprowadzanie różnego rodzaju jednostek. Snow wykazał jej istnienie za-

równy na poziomie komórkowym, jak i poszczególnych osób, gospodarstw domowych, a wreszcie dzielnic.¹⁵

Z tego więc powodu brak odniesienia koncepcji zymiotyczno-miazmatycznej do patologii cholery wydaje się szczególnie istotny. Idea „błądzącego fermentu cholery” w atmosferze i w podobny sposób nieprzewidywalnego skatalizowania przez niego trucizny cholery daje się uzgodnić z danymi tylko na wybranych poziomach analizy, np. grup społecznych czy dzielnic Londynu. Trudno jednak wskazać odpowiednik tej zależności na poziomie jednostkowym czy komórkowym.

Zwróćmy uwagę, że zasadnicze wyjaśnienie, jakie proponowała komisja na gruncie teorii zymiotyczno-miazmatycznej, było *probabilistyczne* na poziomie grup czy dzielnic. Taka bowiem zależność między nasileniem epidemii w poszczególnych dzielnicach a błądzącym fermentem była dopasowana do danych, chociaż ze względu na *probabilistyczny* charakter tego wyjaśnienia dopuszczalna była duża tolerancja niezgodności z danymi przy braku podania powodów tej niezgodności – innych niż zaburzenia losowe.

Przejdźmy na inny poziom analizy: indywidualów czy komórek, wymagałoby dokładnego dookreślenia, od czego zależy prawdopodobieństwo, że wprowadzona do organizmu trucizna cholery rzeczywiście napotka na sprzyjające predyspozycje indywidualum (czy jego organów) do rozwinięcia się choroby. Zgodnie bowiem z tą koncepcją można było wyodrębnić grupę ludzi, którzy zostali poddani zbliżonemu oddziaływaniu tej trucizny i w dalszej kolejności należałoby dookreślić, jakie dokładnie predyspozycje różnią te osoby, które zachorowały, od tych, które nie zachorowały na cholere, mimo poddania oddziaływaniu zbliżonej dawki trucizny.

Podkreślmy w tym miejscu, że w przypadku każdej rozbieżności danych z tym, co przewidywałaby jego koncepcja, Snow był w stanie wytłumaczyć tę rozbieżność, posługując się zaproponowanym mechanizmem rozwoju i przenoszenia cholery.¹⁶ Komisja natomiast pozostawała na przypisaniu tej

¹⁵ Zwróćmy w tym miejscu uwagę, że postępowanie Snowa trudno jest dopasować do wąskiej specjalizacji w nauce. Wiele bowiem z badań, które zrealizował w celu ustalenia zależności między spożywaniem skażonej wody a zachorowalnością na cholere, wykraczało poza ściśle określone kompetencje Snowa jako lekarza (dodajmy: anestezjologa!). Aplikacja proponowanego niżej kryterium niezmienniczości może więc z tytułu daleko bardziej niż w czasach Snowa posuniętej specjalizacji rodzić wiele trudności.

¹⁶ W przypadku koncepcji Snowa mamy do czynienia z określeniem okoliczności, które powodują zmianę funkcjonowania mechanizmu choroby w przypadku, gdy interakcja prowadzi do innych niż oczekiwane rezultatów, np. wyjaśnieniem tego, że w niektórych instytucjach położonych blisko pompy na Broad Street była niska zachorowalność na cholere, jest

losowości *samemu mechanizmowi* oddziaływania, a nie modyfikującym czynnikiem, których nie wskazano.¹⁷

Jak wielokrotnie wykazywał Snow, jedyną okolicznością, która różniła te osoby, było źródło wody. W związku z tym nie byłoby podstaw do sformułowania zależności przyczynowej na poziomie indywiduów na podstawie teorii zymiotyczno-miazmatycznej, gdyż nie byłoby żadnych własności jednostek tego poziomu analizy, do których zwolennik takiej koncepcji mógłby się odwołać.

Być może słuszne jest żądanie w tym miejscu J. Metallmanna, jakie podnosi on za É. Meyersonem:

Natomiast doniosłe jest, że zdarzenia tu występujące są faktami indywidualnymi, zdolnymi do tworzenia klas homogenicznych, tzn. takich, że każdy egzemplarz klasy reprezentuje i zastępuje klasę, i odwrotnie, co dotyczy klasy, dotyczy każdej ze zmian w skład klasy wchodzących; tak że można powiedzieć, iż prawo ogólne *typu przyczynowego* łączy klasy, a przeciw dotyczy indywidualnych zdarzeń, do klas należących. Twierdzenie to kładzie nacisk na *tożsamość* [...] twierdzenie, o którym mowa [...] w istocie pociąga za sobą zasady zupełnie ogólne o charakterze niezmienniczym. (1934, 64)

Niezmienniczość ujawniająca zależności przyczynowe nie wymaga jednak przyjęcia sugerowanych tu mocnych założeń. Ponadto, rozważenie tych założeń – zgodnie z przyjętym we Wprowadzeniu rozróżnieniem – nie jest konieczne w kontynuowaniu rozważań kryterialnych, aczkolwiek może być doniosłe dla filozoficznego namysłu nad przyczynowością.

Powyższe rozważania prowadzą więc do sformułowania niezmienniczości jako kolejnego z warunków adekwatności kryterium zależności przyczynowych.

(N) Jeśli kryterium K jest spełnione w zastosowaniu do danych dla zjawisk X, Y ze zbioru V , to K jest kryterium zależności przyczynowej dla zależności między X a Y , o ile K jest spełnione w zastosowaniu do danych dla dekompozycji V' przeprowadzonej na $\varphi(V)$, gdzie φ oznacza mereologiczną fuzję zjawisk w zbiorze V oraz dla X a Y w

korzystanie z innych źródeł wody. Z kolei wytłumaczeniem tego, że nie wszyscy, którzy pili skażoną wodę, chorowali na cholere, było to, że w przeciwieństwie do trucizn nieorganicznych trucizna cholery nie rozpuszcza się we wszystkich cząsteczkach wody i dlatego nie każda wypita porcja wody jest skażona.

¹⁷ Gruźlicę w XIX wieku – tłumaczono też kwestią predyspozycji, ale wskazywano dokładnie, jakie one były. Tutaj takie wskazanie ani nie zostało podane, ani nie wydawało się możliwe.

zbiorze V istnieją zjawiska X , Y takie, że zachodzi między nimi relacja maksymalnie zwartych części.¹⁸

Spełnienie warunku *niezmienniczości* (N) zależności między X i Y gwarantuje, że zależność między nimi istnieje na różnych poziomach analizy, a więc przy dopuszczeniu różnych jednostek analizy. A co się z tym wiąże, uwzględnienie różnych zmiennych nie zmieni faktu, że X ma wpływ na Y , nawet jeśli na innym poziomie wprowadzono szereg zmiennych pośredniczących, które np. Goldthorpe byłby skłonny interpretować jako mechanizm przyczynowy między tymi zjawiskami. Cox i N. Wermuth podkreślają, że wskazanie zmiennych pośredniczących między X a Y na niższym poziomie analizy niż ten, na którym stwierdzono bezpośrednią zależność przyczynową między tymi zmiennymi, potwierdza, że dokonano właściwej analizy przyczynowej. Wprowadzenie bowiem takich zmiennych pośredniczących przy innej jednostce analizy (na niższym poziomie) jest odpowiednikiem, ich zdaniem, wskazania dokładnego mechanizmu przyczynowego. Jeśli zatem na poziomie grup społecznych Snow wykazał zależność między spożywaniem skażonej wody (X) a zachorowalnością na cholere (Y), to podany przez niego mechanizm rozwoju cholery na poziomie organizmu wprowadza dodatkowe zmienne, które pośredniczą w tej zależności przyczynowej: inkubacja mikroorganizmu skażającego wodę (I), zmiany w naczyniach włoskowatych układu pokarmowego (W), zmiany własności fizyko-chemicznych krwi (K), odwodnienie organizmu (O) i zapaść naczyń krwionośnych w poszczególnych organach (Z).

Inną konsekwencją warunku (N) jest niezmienniczość zależności między X a Y dla różnych badanych populacji i różnych próbek pobranych z tej samej populacji. Jeśli założyć zachodzenie związku przyczynowego na wyjściowym poziomie analizy danych zjawisk, to – w myśl warunku niezmienniczości (N) – mereologiczna fuzja tych zjawisk zachowuje tę zależność i pozwala ją ujawnić także dla innych zjawisk niż w przyjętej wyjściowo dekompozycji, która determinuje jednostkę analizy. Dotyczy to także danych pobieranych z różnych populacji czy różnych próbek z tej samej populacji. W przypadku badań Snowa mówilibyśmy więc np. o tym, że zależność przyczynowa między spożywaniem skażonego pokarmu czy napoju jest nie-

¹⁸ Pojęcie fuzji mereologicznej i jej dystrybucyjnej dekompozycji występuje tu w sensie określonym przez D. Lewisa (1991, 1-3 oraz 22-23). Podobna idea zasygnalizowana – choć nie wyartykułowana – występuje w pracach m.in. Metallmanna (1929/2002) oraz Karpińskiego (1985, 104-105). J. Mączka podkreśla, że w przypadku koncepcji Metallmanna „struktura merologiczna” była uogólnieniem analiz przyczynowych (2001, 145).

zmiennicza dla mieszkańców różnych dzielnic (jako próbek z ogólnej populacji Londynu), a także dla mieszkańców innych miast.

Warto także podkreślić, że warunek niezmienniczości (N) pozwala na uniknięcie pewnej trudności sygnalizowanej przez Metallmanna. Dostrzega on rozbieżność między zależnością „z punktu widzenia *przyczynowego*, która jest zawsze zależnością *klas* zdarzeń, a zarazem zależnością, łączącą *indywidualne* zdarzenia tych klas” a zależnością „ze stanowiska *statystycznego*, która ma stopnie i jest ważna dla zbiorowisk, i *tylko* dla zbiorowisk [...] i to pod warunkiem, że [...] indywidualnie wzięte, są od siebie *niezależne*” (1934, 256). Operacja dekompozycji ujęta w warunku (N) pozwala na powiązanie tych dwóch perspektyw, a w szczególności na uniknięcie problemu międzyjednostkowej przyczynowości, (Spirtes i in. 2000, 295-97) który zapewne był powodem uznania przez Metallmanna tejże rozbieżności.

Zasadnicze pytanie dotyczy tego, czy można wykazać, że warunek (N) jest spełniony na drodze analizy danych? W przypadku bayesowskich modeli zależności przyczynowych taka analiza sprowadza się do sprawdzania warunkowych zależności probabilistycznych przy dwóch założeniach: warunku Markowa oraz warunku wierności. Bardziej szczegółowe omówienie tych warunków pozostawiam do Rozdziału 4. Tutaj zaznaczymy, że warunek Markowa gwarantuje niezależność zmiennych losowych dla poszczególnych przyczyn, a warunek wierności zapewnia, że dla określonej struktury niezależności probabilistycznych istnieje jeden generujący ją mechanizm przyczynowy. Można jednak pokazać, że takie analizy nie gwarantują spełnienia warunku (N) w przypadku ogólnym. Jeżeli bowiem mamy dwie populacje, gdzie zależność probabilistyczna między X a Y wynosi odpowiednio: 0,000 oraz 0,001, to otrzymamy zupełnie różne wnioski co do struktury zależności przyczynowych dla tych populacji. Zatem probabilistyczne zależności warunkowe nie gwarantują zapewnienia, że nie istnieje taka interakcja, która nie zmieni zależności między X i Y , gdyż przy niewielkich już różnicach danych, np. z różnych populacji, testy zależności warunkowych będą prowadziły do całkowicie rozbieżnych wyników.¹⁹

¹⁹ Dla C. Glymoura, opisana przez Coxa i Wermuth procedura wskazywania mechanizmu pośredniczącego między X a Y na niższym poziomie analizy, pełni rolę tylko warunku psychologicznego (rozmowa prywatna). Przy tym podejściu zależność przyczynowa redukuje się do zależności probabilistycznych między zmiennymi przy założeniu warunku Markowa i warunku wierności. Czy wskazanie tego mechanizmu przyczynowego rzeczywiście psychologicznie wzmacnia zależność przyczynową można, zdaniem Glymoura, potwierdzić w serii testów psychologicznych, które sprawdzałyby wykrywalność zależności przyczynowej między dwiema zmiennymi X i Y , a następnie między X , Y i Z , gdzie rozkłady probabilistyczne

Powróćmy więc do badań prowadzonych przez komisję CSI z ramienia the Board of Health. Nawet jeśli zgodzić się, że mechanizm powstawania cholery i nasilania się jej epidemii w poszczególnych lokalizacjach w Londynie ma charakter *wsobnie* probabilistyczny, to i tak warunek (N) wymaga wykazania, że taka probabilistyczna zależność jest niezmiennicza dla różnych populacji i dla różnych poziomów analizy. CSI nie podała żadnych argumentów, które pokazywałyby taką niezmienniczość, a w szczególności nie pokazała, że niezmienniczość ta jest stabilna na niższym, indywidualnym poziomie analizy. CSI nie podała nawet żadnej, choćby przybliżonej charakterystyki tej probabilistycznej zależności, np. dokładnych wartości prawdopodobieństwa przy określonych warunkach atmosferycznych i sanitarnych, ani granic takich wartości. Nawet gdyby jednak przyjąć, że CSI mogła podać taką charakterystykę na prowadzonym poziomie analiz dla grup, to i tak można pokazać, że istnieje interakcja z taką mereologiczną fuzją w obrębie badanej dziedziny rzeczywistości, która niweluje opisaną przez CSI zależność między warunkami atmosferycznymi, „fermentem cholery” a zachorowalnością na tę chorobę. Załóżmy więc, że ten mechanizm przenoszenia cholery, jak chciała CSI, jest *wsobnie* probabilistyczny i że potrafimy podać, choćby zbliżoną, jego charakterystykę względem danych. Potrafimy więc określić, że w pewnych granicach błędu ten mechanizm zadziałał w danych okolicznościach lub nie, a więc że przeniesione z Tamizy przy cieplej pogodzie skażone opary natrafiły na błędzący ferment cholery, że skatalizowana została trucizna, która błędząc natrafiła na określoną lokalizację w Londynie, gdzie została z powietrzem, ewentualnie wodą, wprowadzona do organizmów przebywających tam osób. W znanych jednak przypadkach, jak epidemia przy pompie na Broad Street, istnieją różnice wykraczające poza błąd losowy dopuszczalny w tej koncepcji, różnice, których nie da się wytłumaczyć tą koncepcją ani samym błędem losowym. Warunki okolicznej browarni czy zakładu opiekuńczego nie różniły się niczym od warunków w okolicy Broad Street, a jednak zachorowalność była tam na tyle niższa niż w pobliskiej okolicy, że nie można tego wytłumaczyć tylko błędem losowym. CSI w swojej koncepcji nie uwzględniła zatem jakiegoś istotnego czynnika, który pozwoliłby jej wytłumaczyć ten fakt znany z interakcji z badaną rzeczywistością.

Wydaje się też, że zgodnie z koncepcją CSI konsekwentnie należałoby przyjąć losowość mechanizmu rozwoju choroby na poziomie indywidual-

byłyby takie same, a Z byłoby zmienną pośredniczącą w szczególności, umieszczaną w przestrzeni między X a Y .

nym,²⁰ gdyż wśród osób, które przebywały w tych samych warunkach atmosferycznych tylko niektóre zapadały na cholere. Opracowanie takiej koncepcji mogłoby przebiegać dwojako. Po pierwsze, przez założenie, że ta losowość mechanizmu rozwoju wynika z naszej nieznamomości wszystkich czynników, które są istotne – losowość *epistemiczna*, spowodowana niewiedzą. Gdybyśmy bowiem mieli komplet danych o historii chorób tych osób, ich odporności, czynnikach, na jakie były wystawione itd., moglibyśmy przewidzieć, czy poddanie ich działaniu trucizny cholery spowoduje tę chorobę, czy nie. Po drugie, przez założenie, że ta losowość jest wsobną cechą mechanizmu rozwoju – losowość *ontyczna*, a więc, że niezależnie od kompletności naszej wiedzy na temat osób, które zachorowały na cholere, i czynników, jakim były poddane, wprowadzenie do organizmu trucizny cholery powodowałoby zachorowanie tylko z pewnym prawdopodobieństwem.

Analizy Snowa dotyczące południowych dzielnic Londynu wykazały, że przy losowych czynnikach tak samo działających na populację około trzystu tysięcy osób, jedyną różnicą było źródło wody. Dlatego też to pierwsze wyjaśnienie w oparciu o losowość epistemiczną, a mianowicie, że są nieznanne czynniki, które wyjaśnią dokładnie, dlaczego niektórzy tylko podlegają działaniu błędzącej chmurze trucizny cholery, jest wykluczone w oparciu o inne niż woda czynniki.

Wyjaśnienie natomiast w oparciu o losowość ontyczną prowadzi nieuchronnie do pominięcia konfundera na niższym poziomie analizy. Jeżeli bowiem przyjąć, że X powoduje Y , to – zgodnie z warunkiem niezmienniczości (N) – powinno być możliwe dokonanie dekompozycji mereologicznej fuzji tych zjawisk, w przypadku której zachodziłaby również zależność przyczynowa. Zgodnie jednak z przyjętym założeniem o losowości zależności między X a Y , dokonanie takiej dekompozycji jest niemożliwe do zrealizowania. Oznaczałoby to bowiem, że zależność między X a Y jest wyprowadzalna z innej zależności, a mianowicie z zależności między X' a Y' . Mimo że taka wyprowadzalność nie oznacza jeszcze redukcji zależności między X a Y do zależności między X' a Y' , to pokazywałoby jednak, że losowość tej zależności jest wtórna – i być może – pozorna.²¹

²⁰ Zwróćmy tylko uwagę za E. Eellsem, że na ogół – przy przyjęciu probabilistycznej koncepcji przyczynowości – nie ma związku conceptualnego między różnymi poziomami analiz (Eells 1991, 7-8).

²¹ D. Armstrong podaje uzasadnienie tej tezy na innym gruncie; por. (1985, 29-30).

2.3 Sformułowanie proceduralnego kryterium przyczyny

Studiując postępowanie naukowców we wzorcowych przykładach odkrywania zależności przyczynowych w badaniach nieeksperymentalnych, jak Snow, Doll i Hill, doszedłem do wniosku, że kluczową rolę w dokonanych przez nich odkryciach zależności przyczynowych odegrały własności podjętej przez nich interakcji z badaną rzeczywistością takie, jak zorientowanie na badaną dziedzinę rzeczywistości, ukierunkowanie teoretyczne, a przede wszystkim niezmienniczość zależności przyczynowej niezależnie od wybranego poziomu analizy. Dla zaproponowanego w tym punkcie kryterium proceduralnego zależności przyczynowej istotną rolę odgrywa także to, co decydowało o wyborze takiego, a nie innego poziomu analizy. Zarówno w badaniach Snowa, jak i Dolla, Hilla czy Wyndera,²² dokonywany wybór ukierunkowany był na osiągnięcie odpowiedniego *rodzaju danych*,²³ na podstawie których określano zachodzenie zależności przyczynowej na odpowiednim poziomie analizy. Zmiana poziomu analizy umożliwiła w przypadku badań nad przyczynami cholery, jak i badań nad przyczynami raka płuc, uzyskanie danych uzyskanych dzięki interwencji w badany mechanizm przyczynowy.

O takich danych można powiedzieć, że – na miarę okoliczności – maksymalizują liczbę podjętych przez badacza interakcji z odnośnym fragmentem rzeczywistości.²⁴ Takie dane określać będą poniżej jako *dane interwencyjne*. W przypadku badań eksperymentalnych – przeprowadzonych w idealnych warunkach – przedmiotem interwencji byłoby samo zjawisko X .²⁵ Zgodnie z wprowadzoną poniżej terminologią, w tym przypadku zjawisko X – jako zmienna *manipulowalna* – wyznacza *minimalne pole interwencji* w postaci zależności $X \rightarrow Y$. W przypadku badań nieeksperymentalnych natomiast tylko w niektórych przypadkach można ograniczyć się do analizy danych pochodzących z minimalnego pola interwencji.²⁶

W wielu przypadkach badań nieeksperymentalnych X nie jest zmienną manipulowalną i konieczna jest *ekspansja*. Ta ekspansja może dotyczyć

²² Przegląd dokonań podają R. Doll i współpracownicy (Doll i in. 2004).

²³ Twierdzi się niekiedy, że przełomem w badaniach przyczyn raka płuc było uzyskanie wyników z badań nad bliźniętami jednojajowymi, z których jedno było palaczem, a drugie nie; por. (Kaprio i Koskenvuo 1989).

²⁴ Por. (Harris 2003; Novick i Cheng 2004; Waldmann i Hagmayer 2005).

²⁵ Przegląd trudności, jakie mogą tu wystąpić, podaje A. Sułek (1979, 102-109).

²⁶ Mówiąc w skrócie, chodzi o takie zjawiska, które w danym modelu nie mają przyczyn. Szczegółowo przypadek, gdy $P(X \parallel A) = P(X | A)$, charakteryzuję w Punkcie 4.4.

байдз zmiennych, bądź danych. Jednym z ważnych elementów ustalenia zależności przyczynowej jest przywołanie wyników eksperymentu przeprowadzonego na *innych* zmiennych niż X oraz Y , np. badanie wpływu zależności odkładania się smoły na skórze myszy przez Wyndera i jego współpracowników.²⁷ W takim przypadku mówimy o *ekspansji na zmiennych*. Istotnym elementem takiego rodzaju ekspansji jest zagwarantowanie odpowiedniego związku między polem interwencji $X \rightarrow Y$ a ekspandowanym polem interwencji.

Uzyskanie danych interwencyjnych z minimalnego lub ekspandowanego pola interwencji jest niekiedy niemożliwe lub niecelowe. Wówczas dostępny jest inny rodzaj ekspansji, a mianowicie *ekspansja na danych*. W tym przypadku całkowicie rezygnuje się z poszukiwania danych interwencyjnych, zastępując je *kontrafaktycznymi* danymi interwencyjnymi. Problem ustalenia zależności przyczynowej traktuje się wtedy jako problem brakujących danych, które są dotwarzane za pomocą odpowiedniej analizy dostępnych danych. Z uwagi jednak na to, że w tym przypadku dane nie pochodzą z interakcji badacza z rzeczywistością, lecz są dotworzone za pomocą analizy, ekspansję na danych traktować będziemy jako wtórną wobec ekspansji na zmiennych.

Poniżej proponuję pewne uściślenie wprowadzonych pojęć, przede wszystkim z uwagi na dyskusję w Rozdziale 4, gdzie kryterium proceduralne jest odniesione do najważniejszych współcześnie stosowanych typów modeli przyczynowych: RCM, SEM oraz BN.

Definicja 2.1 (Oddziaływanie czynne)

Działanie A badacza ($A = 1$ lub $A = 0$) oddziałuje czynnie na (zjawisko reprezentowane przez) zmienną X (o stanach x_1, \dots, x_n w przestrzeni probabilistycznej \mathcal{X}) wtw $P(X_\alpha = x_\alpha \mid A = 1, X_{pa(\alpha)}) = P(X_\alpha = x_\alpha \mid A = 1) = 1$ (gdzie $x_i \in \mathcal{X}_i \subset \mathcal{X}$ dla $1 \leq i \leq n$).²⁸

Dla uproszczenia przyjmujemy, że jedno działanie A oddziałuje czynnie na co najwyżej jedną zmienną w danej dziedzinie D .

²⁷ Zgodnie z warunkiem niezmienniczości (N) ekspansja na zmiennych jest szczególnym przypadkiem przeprowadzenia dekompozycji mereologicznej fuzji zjawisk, które reprezentuje zmienna X .

²⁸ Wprowadzoną tu terminologię szczegółowiej objaśniam w Punkcie 4.4.

Definicja 2.2 (Oddziaływanie bierne)

Działanie A badacza ($A = 1$ lub $A = 0$) oddziałuje biernie na zmienną X (o stanach x_1, \dots, x_n w przestrzeni probabilistycznej \mathcal{X}) wtw $P(X_\alpha = x_\alpha | A = 1) \neq P(X_\alpha = x_\alpha | A = 0, X_{pa(\alpha)}) \neq P(X_\alpha = x_\alpha | X_{pa(\alpha)}) = 1$.

Oddziaływanie bierne wprowadza więc pewną zmianę w istniejącym układzie, jednak staje się jednym z wielu czynników oddziałujących na badane zjawisko.

Działanie A badacza *oddziałuje* na zmienną X wtw A oddziałuje na X czynnie lub biernie.

Definicja 2.3 (Zmienna manipulowalna)

Zmienna X , dla której istnieje działanie A spełniające definicję oddziaływania czynnego na X , jest zmienną manipulowalną.

Zmienna X , która nie jest manipulowalna, jest zmienną obserwacyjną.

Jeśli dane są pobierane, gdy X jest poddana oddziaływaniu czynnemu ($A = 1$), mówimy o *danych interwencyjnych*.²⁹

Definicja 2.4 (Dane interwencyjne)

Zbiór $\mathbf{I}_X^V = \times_{i \in X} \mathbf{V}_i \times_{j \in V \setminus X} \mathbf{V}_j$ wektorów wartości zmiennych $x_i \in X \subset V$ oraz $v_j \in (V \setminus X) \subset V$ z dziedziny D pobranych z próbki lub populacji, gdy A oddziałuje czynnie na zmienne w X , określamy jako dane interwencyjne X .

Tak więc interwencja na zjawiska reprezentowane przez zmienne w zbiorze X „przecina” zbiór danych na dwa podzbiory: zbiór danych *czysto interwencyjnych* dla zmiennych w X poddanych manipulacji oraz zbiór danych dla pozostałych zmiennych. Będziemy też posługiwać się skróconym oznaczeniem \mathbf{I}_X .

Istnieją także przypadki, gdy można mówić o danych interwencyjnych mimo braku działania badacza (człowieka) na dane zjawisko, np. w przypadku eksperymentu naturalnego z dystrybucją wody w południowych dzielnicach Londynu wskazanego przez Snowa.

Wartości zmiennej $X = x_i$ pobrane, gdy A nie oddziałuje na X , określamy jako *dane obserwacyjne*.

²⁹ Termin „interwencja” jest na ogół traktowany zamiennie z terminem „interwencja”; por. (Sulek 1979). Wyjątkiem jest koncepcja Woodwarda, gdzie pojęcie (ludzkiej) interwencji traktuje się jako szczególny przypadek ogólniejszego pojęcia manipulacji.

Definicja 2.5 (Pole interwencji)

Polem interwencji w dziedzinie D dla zmiennej Y jest taki zbiór $PI = (\langle X_1, \dots, X_n \rangle, Y)$, że dla zbioru zmiennych $(X_1, \dots, X_n) \subseteq X$ o danych interwencyjnych \mathbf{I}_X^V zachodzi $f(Y | V) = f(Y | (X_1, \dots, X_n))$, gdzie $f(X_i)$ jest funkcją gęstości rozkładu.³⁰

Sens tej definicji jest następujący. Załóżmy, że spośród zjawisk V wybieramy pewne, a mianowicie X , i poddajemy je manipulacji. Stwierdzamy następnie, że dla wyjaśnienia zmian wartości Y wystarczy uwzględnić część spośród zjawisk w X , a mianowicie X_1, \dots, X_n . W przypadku stwierdzenia, że oprócz zjawisk poddanych manipulacji w badaniu występują inne jeszcze, które mogą mieć wpływ na Y , należy ponownie opracować plan badania, który uwzględni także tamte zmienne.

Istotne pytanie dotyczy estymacji pola interwencji przed opracowaniem planu badania.³¹ Szczególnym przypadkiem jest sytuacja, gdy manipulowana jest tylko jedna zmienna. Otrzymujemy wówczas zależność $f(Y | V) = f(Y | X_i)$. Minimalnym polem interwencji dla zmiennej Y wyznaczonym przez zmienną X jest więc pole interwencji $PI = (\langle X, Y \rangle)$.

Jednym z wniosków, jakie płyną z tej definicji, jest to, że w przypadku zmiennych X i Y , gdzie X jest zmienną obserwowalną, nie istnieje minimalne pole interwencji.

Kolejny wniosek z podanej definicji jest taki, że w przypadku, gdy X jest zmienną manipulowalną, zależność $X \rightarrow Y$ jest niezmiennicza pomiędzy polem interwencji a minimalnym polem interwencji.

Jak podkreślałem wcześniej, zasadniczym celem argumentacji Snowa było nie to, aby wykazać, że skażona woda jest *jedną* z przyczyn cholery³², ale że jest ona *zasadniczą*, czyli aktywną, przyczyną tej choroby. Najprostszym sposobem wykazania tego, jak można byłoby sądzić, polegałby bądź 1) na podawaniu takiej wody losowo wybranym osobom i obserwowaniu skutków, bądź 2) na poddaniu tej wody szczegółowej analizie i zaobserwowaniu postulowanego przez Snowa żywego organizmu, który byłby w niej obecny. Możliwość pierwsza, z uwagi na nieznaną mechanizm choroby i brak środków zapobiegawczych, była niedostępna ówczesnie. Jest możliwa dziś,³³

³⁰ W zasadzie jest to pojęcie pokrycia Markowa, tyle że ograniczone do przypadku danych interwencyjnych.

³¹ Por. (Brzeziński 1996, 282-342).

³² W sensie przyczyny dopełniającej lub przysposabiającej.

³³ O współczesnych badaniach tego rodzaju piszą w renomowanym czasopiśmie *Lancet* D. Sack i in. (2004).

ale warunkiem tego jest wcześniejsze poznanie mechanizmu cholery, a celem dokładniejsze poznanie mechanizmu, np. jaka ilość bakterii wystarcza do wywołania choroby u osób o różnym poziomie produkcji soku żołądkowego czy różnych grupach krwi. Możliwość druga – jak się wydaje – również była niedostępna – przynajmniej dla Snowa – ale tylko ze względów technicznych. Powszechnie przyjmuje się też, że zaobserwowanie *vibrio cholerae* przez Kocha było decydującym krokiem w poznaniu cholery (Howard-Jones 1984).

Jeżeli zatem rozważamy *ekspansję* pola interwencji „Skażona woda → Cholera”, to w świetle powyżej poczynionych uwag oraz kryterium niezmienniczości (N) jasne staje się, że dokonuje się ona poprzez zmianę jednostki analizy zjawisk „skażona woda” oraz „cholera”. Z uwagi na problemy międzyjednostkowej przyczynowości (Spirtes i in. 2000, 295-97) oraz międzygatunkowej przyczynowości (Freedman i Zeisel 1988a; 1988b) zjawiska będące (mereologicznymi) składowymi zjawisk pola interwencji „Skażona woda → Cholera” nie mogą stać się składowymi tegoż pola interwencji. Mereologiczna fuzja zjawisk składających się na wspomniane pole interwencji pozwala na jej dekompozycję, która wyodrębni nowe jednostki analizy. Taką procedurę określać będę niżej jako *ekspansję* tego pola interwencji. W przypadku pola interwencji „Skażona woda → Cholera” jednostką analizy jest pojedyncza osoba. Mereologiczną fuzję wszystkich osób mieszkających w południowych dzielnicach Londynu lub korzystających z usług dostawców wody w tamte rejony miasta można dekomponować na różne rodzaje zbiorów zjawisk, odpowiadające różnym jednostkom analizy. Można więc tej dekompozycji dokonać, wprowadzając relację maksymalnie zwartych części w stosunku do komórek – zwłaszcza krwi i nabłonka kosmków jelitowych – i ich własności fizyko-chemicznych oraz w stosunku do grup mieszkańców, wyodrębnionych z uwagi na różnych dostawców wody.

Zgodnie z warunkiem niezmienniczości (N), jeśli związek przyczynowy zachodzi na poziomie jednostkowym, to także na poziomach wyróżnionych na drodze opisanej wyżej dekompozycji.³⁴ Zjawisko „skażona woda” odnosi się do rodzaju wody, której sposób filtrowania nie gwarantował usunięcia „aktywnego czynnika” powodującego, zdaniem Snowa, cholere. Nie można na tym poziomie analizy jednoznacznie stwierdzić, że „skażona woda” zawiera bakterię wywołującą cholere. Jednak niemal jednoznacznie można to stwierdzić o „nieskażonej wodzie”, która dostarczana była mieszkańcom

³⁴ Przypomnę za Lewisem, że dekompozycję wyznacza dobór relacji bycia maksymalnie zwartą częścią (1991, 22-23).

południowych dzielnic Londynu. Snow przyjmował jednak, że w przypadku każdego zachorowania na cholereę można wyjaśnić to, odwołując się do wprowadzenia do organizmu niewidzialnego mikroorganizmu wraz z pokarmem lub wodą. Jednostkowy poziom rozważań nie był według Snowa odpowiedni do uzasadnienia mechanizmu działania choroby z uwagi na niemożliwość wykrycia obecności owego pasożytniczego mikroorganizmu ani dokonania na tym poziomie eksperymentu, który tę obecność mógłby pośrednio udowodnić. W sposób pośredni jego obecność i efekty aktywności w organizmie gospodarza można było wykryć na poziomach innych niż jednostkowy, np. komórek lub grup.

Przyjęta przez Snowa strategia ekspansji pola interwencji przebiegała w jednym z dwóch kierunków: bądź uzyskania pomiarów na próbie statystycznej, pozwalających na określenie schematów zależności probabilistycznych dla populacji (poziom grup), bądź uzyskania pomiarów na niewielkiej próbie, którą charakteryzuje duża jednorodność przyczynowo-skutkowa (poziom komórek).

Przy ekspansji pola interwencji we wnioskowaniu Snowa poziom jednostkowy odgrywa rolę integrowania informacji uzyskanych w bezpośredniej interakcji z rzeczywistością. Dla każdego przypadku osoby zmarłej na chorobę – poprzez poziom jednostkowy – postuluje się przyczynę w postaci wprowadzenia do organizmu bakterii wraz z piciem lub jedzeniem. W ten sposób formułuje się mechanizm wyjaśniający dane oraz stwierdzone w nich zależności statystyczne między spożywaniem skażonej wody a śmiertelnością wywołaną cholera.

Podobnie dla poziomu mikro. Stwierdzenie obecności bakterii *vibrio cholerae* w pobliżu komórek jelita, zwłaszcza kosmków jelitowych, znowu poprzez poziom jednostkowy, tłumaczy się wprowadzeniem do organizmu nosiciela bakterii wraz z wodą lub pokarmem.

Za rozstrzygające dowody podane przez Snowa uznaje się: analizę statystyczną naturalnego eksperymentu ze spożyciem skażonej wody w Londynie oraz efekt zdjęcia ramienia z pompy przy Broad Street.³⁵ W pierwszym przypadku w dwóch naturalnie utworzonych losowych grupach mierzono

³⁵ Jak twierdzi N. Paneth (2004, 514), historyczny fakt zdjęcia ramienia pompy przy Broad Street na polecenie Snowa należy poddać w wątpliwość. Faktycznie, usunięte zostało ramię innej pompy. Podkreślimy jednak, że wbrew powszechnemu przekonaniu Snow nie łączył spadku liczby zachorowań w tym rejonie Londynu ze zdjęciem ramienia pompy.

śmiertelność spowodowaną cholera oraz dostawcą wody.³⁶ W przypadku pompy znana była śmiertelność mieszkańców okolicznych dzielnic Londynu spowodowana cholera przed zdjęciem ramienia i po nim, chociaż zmiana populacji tych dzielnic spowodowana nagłą emigracją nie była rejestrowana.

Snow uznawał, że te zjawiska z poziomu populacji dają się wyjaśnić na poziomie jednostkowym. Dla każdego przypadku zachorowania na cholera postulował przyczynę w postaci wprowadzenia do układu trawiennego mikroorganizmu powodującego gwałtowne odwodnienie, najczęściej wraz z wypitą wodą, ale także pokarmem. Przyjmował też implicite założenie, że przebieg choroby na poziomie mikro jest jednorodny.

To założenie o roli poziomu jednostkowego w stwierdzaniu zależności przyczynowej jest wyeksponowane w postępowaniu zalecanym przez tzw. postulatory Kocha.³⁷ Żąda się w nich, aby stwierdzenie tej zależności na poziomie populacji oraz na poziomie mikro było każdorazowo zapośredniczone przez poziom jednostkowy (wprowadzanie bakterii lub jej izolowanie z organizmu danego zwierzęcia).

Podsumowuję powyższe rozważania, proponując uściślenie pojęcia ekspansji pola interwencji $X \rightarrow Y$, a następnie kryterium proceduralne.

Definicja 2.6 (Ekspansja minimalnego pola interwencji)

Ekspansją minimalnego pola interwencji $PI = \langle X, Y \rangle$ jest pole interwencji $PI' = \langle X', Y' \rangle$ takie, że X', Y' są elementami dekompozycji fuzji mereologicznej zbioru zjawisk V (do którego należą X oraz Y) oraz $R(X', X)$, oraz $R(Y', Y)$, gdzie R jest relacją bycia maksymalnie zwartą częścią lub jej konwersem oraz istnieją dane interwencyjne $I_{X'}$ dla zbioru X' .

Definicja 2.7 (Kryterium proceduralne przyczyny)

Zależność $X \rightarrow Y$ jest przyczynowa w dziedzinie D wtw dla danych interwencyjnych I_X^V w minimalnym polu interwencji Y wyznaczonym przez X i wartościach zmiennych w V z dziedziny D zachodzi $f(Y | V) = f(Y | X)$ lub dla danych interwencyjnych X' w ekspansji minimalnego pola interwencji

³⁶ Jakość wody ulegała wahaniom, jednak różnica stosowanych metod filtrowania dawała w efekcie stałą różnicę jakości wody.

³⁷ 1. Dany mikroorganizm powinien być obecny u wszystkich zwierząt chorych, a nieobecny u zdrowych. 2. Powinno się go wyizolować z chorego zwierzęcia i wyhodować w czystej kulturze w warunkach laboratoryjnych. 3. Po wprowadzeniu tego wyizolowanego mikroorganizmu do zdrowego zwierzęcia, powinno się zaobserwować tę samą chorobę, co u wyjściowego. 4. Powinno się go dać ponownie wyizolować (znaleźć) u zainfekowanego eksperymentalnie zwierzęcia.

Y' wyznaczonej przez X' zachodzi $f(Y' | X', V) = f(Y' | X')$, gdy X nie jest zmienną manipulowalną (V zbiorem jest zmiennych o wartościach w D).

X jest traktowana jako manipulowalna także wtedy, gdy $P(Y || X) = P(Y | X)$.

W ogólnej postaci model zależności zmiennych to $M = M(V, R)$, gdzie V jest zbiorem zmiennych uwzględnionych w modelu, a R zbiorem zależności między zmiennymi w V .

Definicja 2.8 (Kryterium proceduralne przyczyny względem modelu)

Zależność $X \rightarrow Y$ jest przyczynowa w modelu $M(V, R)$ wtw dla danych nieinterwencyjnych X w minimalnym polu interwencji Y wyznaczonym przez X i o wartościach zmiennych V w D zachodzi $f(Y | V) = f(Y | X)$ lub dla danych nieinterwencyjnych X' w ekspansji minimalnego pola interwencji Y' wyznaczonej przez X' zachodzi $f(Y' | X', V) = f(Y' | X')$, gdy X nie jest zmienną manipulowalną (V jest zbiorem zmiennych o wartościach w D).

Przyjęta w zaproponowanym wyżej kryterium proceduralnym precyzacja warunku niezmienniczości (N) jest ukierunkowana przez dążność do osiągnięcia danych interwencyjnych.³⁸ Niezmienniczość zależności przyczynowej nie określa, jak sędzę, w wystarczającym stopniu, jaki należy wybrać poziom analizy, aby stwierdzić zachodzenie zależności przyczynowej.³⁹ Takie ukierunkowanie natomiast w naturalny sposób wynika z dążności do pozyskania danych interwencyjnych dla dokonywanej dekompozycji. Z warunku (N) wynika, że w przypadku, gdyby istniało wiele takich dekompozycji umożliwiających pozyskanie danych interwencyjnych, wówczas o wyborze jednej z nich mogą decydować względy optymalizacji badań z uwagi na np. dostępność środków technicznych, kosztowność, czasochłonność, itp.

³⁸ Alternatywną, powiedziałbym też bardziej tradycyjną, eksplikację warunku (N) dyskutował T. Pawłowski (206-26).

³⁹ Z innej perspektywy ten sam problem opisuje Lewis, zwracając uwagę, że mereologiczna fuzja tylko w wyjątkowych przypadkach determinuje jednoznacznie dopuszczalną dekompozycję do postaci zbioru w sensie dystrybutywnym, co uzależnione jest przede wszystkim od tego, jakie zostaną wyodrębnione maksymalnie zwarte części tej fuzji; por. (1991, 22-23).

3

Kryterium proceduralne jako uogólnienie jakościowych kryteriów przyczyny

3.1 B. Hilla kryteria przyczyny

Odpowiedzi na pytanie o to, jak zaproponowane w Punkcie 2.3 kryterium proceduralne odnosi się do kryteriów zależności przyczynowych dyskutowanych w literaturze metodologicznej, należy szukać przede wszystkim w konfrontacji z kryteriami B. Hilla (1965).¹ W literaturze metodologicznej poświęconej odkrywaniu zależności przyczynowych między badanymi zjawiskami w badaniach nieeksperymentalnych są to najczęściej przywoływane kryteria. Często wykorzystywane są one do interpretacji otrzymanych wyników w badaniach naukowych. W dyskusji uwzględniłam również zmodyfikowane ich wersje (Susser 1973; Cox i Wermuth 1996) oraz ich krytykę (Rothman, Greenland 1998).

Kryteria, jakie opracował Hill, są rozszerzoną wersją kryteriów zaproponowanych w raporcie U.S. Surgeon General's Advisory Committee *Smoking and health* (1964). Antycypowali je² J. S. Mill, podając metody indukcji w *Systemie logiki* (1843/1962), oraz D. Hume, podając swoje reguły w *Badaniach* (1739/1963).

Jak podkreśla Hill, te kryteria mają pomóc w określeniu, kiedy związek między dwoma zjawiskami można zinterpretować przyczynowo. Jako pierwsze z kryteriów wymienia on *siłę* asocjacji między zmiennymi. Powołując się na wyniki własnych kilkunastoletnich badań nad przyczynami raka

¹ W literaturze polskiej omawiają je A. Zieliński i J. Mazurek (2002, 280-86).

² Por. (Rothman, Greenland 1998, 24).

pluc, Hill podaje, że śmiertelność z tego powodu wśród palaczy papierosów jest dziewięć do dziesięciu razy większa niż u niepalących, a u tych, którzy palą bardzo dużo, ta proporcja wzrasta dwu- lub trzykrotnie. Natomiast śmiertelność wśród palaczy z powodu zakrzepów w naczyniach wieńcowych jest nie więcej niż dwukrotnie wyższa w porównaniu z tymi, którzy nie palą papierosów. W tym drugim przypadku można podać interpretację przyczynową, ale – jak sugeruje Hill – można również podejrzewać, że istnieje pewne zjawisko związane z paleniem papierosów, które jest przyczyną obu zjawisk, np. brak ruchu czy nieodpowiednia dieta. W przypadku raka natomiast

wyjaśnienie stwierdzonego wzrostu raka płuc w kategoriach innych własności środowiskowych wymaga przywołania takiej własności życia, która byłaby tak niezwykle ściśle związana z paleniem papierosów oraz z jego natężeniem, że dawałaby się łatwo wykryć. (1965, 295)

Gdyby natomiast ktoś twierdził, że może jest taka własność, ale nie daje się jej wykryć ani nawet dokładniej określić, to „mamy prawo odrzucić niejasną krytykę tego teoretyka” (1965, 296). Jako przykład Hill podaje także badania Snowa podjęte latem 1854 roku, które wskazywały czterynastokrotnie wyższą zachorowalność na cholera w ciągu czterech pierwszych tygodni epidemii wśród tych, którzy korzystali z wody dostarczanej przez Southwark i Vauxhall w porównaniu z Lambeth.³

Drugim na liście kryteriów Hilla jest *spójność* zaobserwowanej asocjacji: „Czy powtarzała się w obserwacjach dokonywanych przez różne osoby, w różnych miejscach, okolicznościach i czasie?” (1965, 297).⁴ Przykładem takiej spójności są retrospektywne i prospektywne badania nad rakiem płuc, które – mimo różnych okoliczności i technik – prowadziły do tej samej odpowiedzi. Jednak to kryterium nie jest konieczne dla zależności przyczynowej, co Hill ilustruje następującym przykładem. Pacjenci oczekujący w szpitalu na operację z powodu wrzodu układu trawiennego odpowiadają na pytanie o niedawne kryzysy czy kłótnie domowe. Podobnie pacjenci oczekujący na operację przepukliny. W tym pierwszym przypadku zgłoszenie się do szpitala z pękniętym wrzodem jest natychmiastowe, w drugim natomiast – może być oddalone w czasie. To zaś może prowadzić do systematycznego

³ Jednak, jak sugeruje Hill, siła asocjacji nie jest warunkiem koniecznym przyczynowości, gdyż znane są w medycynie przypadki, gdy zależność przyczynowa jest słaba, np. niewiele osób, które podczas pracy narażone są na kontakt z moczem szczurów, choruje na chorobę Weilla.

⁴ Zieliński i Mazurek (2002, 281) używają tu terminu „powtarzalność”.

błędu w wyciąganiu wniosku o zależności przyczynowej między kryzysem domowym a przepukliną, którego nie wykryje żadne systematyczne powtarzanie obserwacji.

W przypadku spójności, jak zwraca uwagę Hill,

Mamy zatem nieco paradoksalne stanowisko, że różne wyniki różnych badań nie mogą być podstawą odrzucenia pierwotnych dowodów empirycznych; a jednak te same wyniki uzyskane w tego samego rodzaju badaniach nie będą na ogół wzmacniać pierwotnych dowodów empirycznych. Osobiście wielką wagę przywiązywałbym do osiągnięcia podobnych rezultatów różnymi sposobami, np. prospektywnie i retrospektywnie. (1965, 296)

W przypadku więc, gdy mamy do czynienia z jedną tylko sytuacją, której nie można powtórzyć, lecz związek między zjawiskami jest bardzo mocny (np. śmiertelność z powodu raka nosa i płuc wzrasta ponad stukrotnie), nie mamy wątpliwości co do interpretacji przyczynowej mimo braku powtarzalności, a co za tym idzie – spójności.

Trzecim diskutowanym kryterium jest *specyficzność*.⁵ Jeśli śmiertelność jest związana z bardzo określonymi zachorowaniami (choroby układu oddechowego) i dotyczy pewnej grupy pracowników jakiegoś zakładu w określonym miejscu, należy uznać to – zdaniem Hilla – za mocny argument na rzecz interpretacji przyczynowej tej zależności. Nie jest to jednak warunek konieczny takiej interpretacji, gdyż jedna przyczyna może mieć więcej niż jeden skutek i wówczas warunek specyficzności nie może być spełniony.

W przypadku palenia papierosów odnotowuje się wzrost śmiertelności nie tylko z powodu raka płuc, ale i innych chorób. Jednak w tym przypadku, jak podkreśla Hill, specyficzność przejawia się w sile związku: śmiertelność z powodu raka płuc wzrasta o 900%-1000%, natomiast dla innych chorób jest to 10%-50%. Zatem, zdaniem Hilla, specyficzność, mimo że rzadko obserwowana (ze względu na częstą wieloczynnikową przyczynowość), jest w zasadzie warunkiem wystarczającym interpretacji przyczynowej.

Czwartym kryterium jest *następstwo* czasowe. Dotyczy ono przede wszystkim takich sytuacji, gdy efekt lub przyczyna mogą być rozciągnięte w czasie. Przykładowe pytania, jakie należy tu postawić, to: czy dana dieta doprowadziła do tej oto choroby, czy też wczesne etapy tej choroby spowodowały wykształcenie takich a nie innych nawyków żywieniowych? Czy zarażenie gruźlicą jest związane z miejscem pracy, czy też taki, a nie inny

⁵ Zieliński i Mazurek (2002, 281) posługują się dla tego kryterium terminem „swoistość”.

rodzaj zajęcia wybierają osoby skłonne do zachorowania na gruźlicę lub nią już zakażone?

Piątym kryterium jest *gradient biologiczny*. Dane, które ujawniają gradient biologiczny czy krzywą dawka-odpowieź, w dużym stopniu skłaniają do interpretacji przyczynowej. Zdaniem Hilla, bardzo wyraźnym tego przykładem jest liniowa zależność między ilością palonych papierosów a wzrastającą śmiertelnością z powodu raka płuc.

Na szóstym miejscu listy Hilla znalazła się *wiarygodność*. Nie jest warunkiem koniecznym, ale pomocnym wiedzieć, że dana asocjacja jest prawdopodobna biologicznie. Ocena, czy spełnione jest to kryterium, zależy od obecnej wiedzy biologicznej. Hill podkreśla jednak, że mogą być przypadki takich asocjacji, dla których nie posiadamy na obecnym etapie rozwoju nauki wystarczającej wiedzy, a mimo to może ona być odzwierciedleniem zależności przyczynowej.

Siódmym kryterium jest *koherencja*. Hill zaznacza, że interpretacja przyczynowa danej asocjacji nie może być niezgodna z ogólną wiedzą przyrodniczą oraz wiedzą biologiczną o danej chorobie. I tak the Advisory Committee to the Surgeon General uznało interpretację przyczynową asocjacji między paleniem papierosów a rakiem płuc za koherentną z wiedzą o nasileniu się obu tych zjawisk w latach 1950. oraz z różnicą śmiertelności dla obu płci.

Hill podkreśla, że dla niego osobiście koherencję zależności między paleniem papierosów a rakiem płuc znacznie wzmocniły dowody empiryczne z badań nabłonka oskrzeli palaczy oraz wyizolowanie z dymu papierosowego czynników rakotwórczych dla skóry zwierząt laboratoryjnych. Hill zwraca jednak uwagę, że

podczas gdy takie laboratoryjne dowody empiryczne mogą znacząco wzmocnić hipotezę oraz określić czynniki przyczynowe, brak takich dowodów nie może zniweczyć epidemiologicznych asocjacji w człowieku. (1965, 297)

A następnie odnosi ten komentarz do badań prowadzonych przez Snowa:

W szerszym kontekście obserwacje epidemiologiczne Johna Snowa o przeniesieniu cholery za pośrednictwem wody z pompy przy Broad Street byłyby poza wszelką wątpliwością, gdyby na miejscu był Robert Koch, który wyizolowałby vibrio z pieluszek dziecka, studni głębinowej i z organizmu mężczyzny o delikatnym zdrowiu z Brighton. Jednak fakt, że na prace Kocha trzeba było czekać następnymi trzydziści lat, nie osłabiał siły tego epidemiologicznego argumentu, chociaż utrudnił jego obronę przed ówczesnymi krytykami – zarówno słusznymi, jak i niesłusznymi. (1965, 298)

Ósmym kryterium Hilla jest *eksperyment*.⁶ Niekiedy dostępne są dowody empiryczne pochodzące z eksperymentu lub quasi-eksperymentu. Niekiedy też w oparciu o stwierdzoną asocjację dokonuje się akcji prewencyjnej i można postawić pytanie, zdaniem Hilla, o to, czy takie działanie rzeczywiście przynosi żądany skutek, np. redukcja zapylenia w miejscu pracy, zmiana rodzaju oleju do smarowania maszyn czy zaprzestanie palenia papierosów. Hill stawia więc pytanie: „Czy to zmienia częstość skojarzonych zdarzeń? Tutaj może ujawnić się najmocniejsze potwierdzenie hipotezy przyczynowej” (1965, 297).

W niektórych okolicznościach dopuszczalne dla Hilla jest także stosowanie *analogii*, co zamyka listę dziewięciu kryteriów interpretacji przyczynowej asocjacji między zmiennymi. W podsumowującym komentarzu Hill stanowczo odmawia tym kryteriom charakteru automatyczności:

Nie wierzę – jak to już zostało zaznaczone – że możemy w pożyteczny sposób podać żelazne reguły nakładane na dowody empiryczne, które muszą być przestrzegane przed uznaniem zależności przyczynowo-skutkowej. Żadne z moich dziewięciu punktów widzenia nie może dostarczyć bezdyskusyjnego dowodu za lub przeciw hipotezie przyczynowej i żaden nie może być wymagany jako *sine qua non*. Mogą one pomóc nam, w mniejszym lub większym stopniu, podjąć decyzję w odniesieniu do fundamentalnego pytania: Czy jest jakiś inny sposób wytłumaczenia zestawu faktów, jaki mamy przed sobą? Czy jest jakiś inny sposób, bardziej prawdopodobny niż zależność przyczynowo-skutkowa? (1965, 298)

W konsekwencji więc, zdaniem Hilla, nie ma takiego *formalnego* testu zależności przyczynowej między zmiennymi.⁷ Testy formalne co najwyżej

mogą i powinny przypominać nam o efektach, jakie gra przypadku może powodować i pozwolą nam ocenić prawdopodobną wielkość tych efektów. Poza tym jednak w żaden sposób nie przyczyniają się do podania „dowodu” naszej hipotezy. (1965, 300)

⁶ Trafniejsze wydaje się mniej literalne tłumaczenie jako „potwierdzenie w badaniach eksperymentalnych” (Zieliński i Mazurek 2002, 284).

⁷ Hill w tekście przywołuje test istotności, który był powszechnie wówczas stosowaną techniką statystyczną. Jego argument odnosi się jednak szerzej do wszelkich metod formalnych (czysto analitycznych).

3.2 M. Sussera kryteria przyczyny

Cox i Wermuth (1996; 2001) modyfikują listę kryteriów Hilla następująco. Przede wszystkim odnoszą je do sytuacji, gdy dla zaproponowanego modelu zależności przyczynowej między zmiennymi można określić mechanizm generujący tę zależność, najczęściej na niższym poziomie analizy. Jako pierwsze dwa warunki wymieniają zgodność teoretyczną zaproponowanej interpretacji przyczynowej:

1. jest bardziej prawdopodobne, że efekt jest przyczynowy, jeśli wyjaśnienie jest mocno ugruntowane w wiedzy przedmiotowej tego efektu oraz został wcześniej przewidziany;
2. podobne wyjaśnienie podane retrospektywnie zwiększa możliwość, że dany efekt jest przyczynowy, ale w mniej przekonujący sposób niż wtedy, gdy jest podane wcześniej. (1996, 225)

Te dwa warunki łączą zaproponowane przez Hilla standardy prawdopodobieństwa i koherencji, które znalazły się odpowiednio na szóstym i siódmym miejscu jego listy.

Dalsze warunki, jakie podają Cox i Wermuth – używając określeń Hilla – są następujące: siła (jest bardziej prawdopodobne, że większe efekty mają interpretację przyczynową niż małe), gradient biologiczny (jest bardziej prawdopodobne, że efekty monotoniczne przy użyciu pewnej naturalnej miary „dawki” są przyczynowe, niż przy efektach nieregularnie zachowujących się wobec dawki, chociaż w niektórych przypadkach należy oczekiwać gładkiego niemonotonicznego zachowania), spójność (występowanie efektu w niezależnych badaniach, zwłaszcza o nieco różnej formie, zwiększa prawdopodobieństwo interpretacji przyczynowej), specyficzność przyczyn (bardziej prawdopodobna jest interpretacja przyczynowa skutków, które „nie wykazują większych interakcji z wieloma podstawowymi (ang. intrinsic) zmiennymi wyjaśniającymi”), eksperyment naturalny („masowe interwencje” mogą dostarczać informacji o zależnościach przyczynowych) i specyficzność natury skutków (Cox i Wermuth 1996, 225).

Te modyfikacje listy Hilla odzwierciedlają przede wszystkim przekonanie Coxa i Wermuth, że do sformułowania interpretacji przyczynowej modelu danych konieczne jest oparte na wiedzy przedmiotowej rozumienie procesu generującego te dane. Zróżnicowanie na wyjaśnienie podane przed i po wyniku, zdaniem tych autorów, z faktu, że „W wielu dziedzinach możliwe jest przy odrobinie wysiłku podanie prawdopodobnego wyjaśnienia prawie dowolnego wyniku obserwacji” (1996, 226).

W komentarzu do kontrowersyjnego warunku specyficzności Cox i Wer-muth dodają, że odgrywa on istotną rolę w przypadku, gdy mamy do czynienia z dobrze określonym fizycznie lub biologicznie procesem, natomiast zdecydowanie mniejszą wówczas, gdy brane pod uwagę zmienne są w rzeczywistości konglomeratami różnych własności (1996, 226-27).

Jeszcze dalej idącą modyfikację standardów Hilla znajdujemy w klasycznej pracy Sussera *Causal thinking in the health sciences* (1973). W charakterystyczny dla siebie sposób Susser upatruje zasadniczej trudności w przyczynowej interpretacji zależności między zmiennymi w „złożonym powiązaniu wszystkich poziomów organizacji w ekologicznych systemach realnego świata” (1973, 72). Dlatego kryteria Hilla stanowią jeden z elementów strategii „radzenia sobie ze zmiennością biologicznych i społecznych zjawisk w populacjach oraz z niepewnością dotyczących ich wnioskowań przyczynowych”. Owe elementy obejmują:

Po pierwsze, uproszczenie warunków obserwacji przez wyizolowanie i kontrolowanie danego segmentu rzeczywistości; po drugie, przesianie modeli pod względem źródeł zmienności zewnętrznej lub nierelevantnej w stosunku do modelu; po trzecie, przetestowanie zaobserwowanych asocjacji przez opracowanie ich związków w analizie; po czwarte, wykorzystanie pojęć prawdopodobieństwa do oszacowania istotności związanej z współwystępowaniem zmiennych w obserwowanych przypadkach; po piątę, poddanie zebranych danych pod osąd kilku określonych kryteriów. (Susser 1973, 72)

Mówiąc o uproszczeniu warunków obserwacji, Susser podkreśla, że wyodrębnienie badanego wycinka rzeczywistości zależy od danej dyscypliny, niekiedy przypadkowych elementów związanych z jej rozwojem, ale także od ograniczeń świadomie wprowadzonych przez badacza. Planując badanie, ogranicza się istotny wycinek rzeczywistości z uwagi na postawione pytanie oraz możliwości zebrania danych, które mają służyć odpowiedzi na to pytanie.⁸ Jak podkreśla Susser, „Izolacja systemu nigdy nie jest doskonała, czynniki zewnętrzne oddziałują na relacje między przyczyną a skutkiem” (1973, 73).

W przypadku badań nieeksperymentalnych wykluczona jest randomizacja, która ma na celu porównanie działania zmiennej niezależnej na zależną w dwóch populacjach, które różni jedynie to właśnie działanie. Susser wskazuje inne sposoby planowania badań: badań kohortowych, w których porównuje się wystarczająco liczną grupę osób wystawioną na działanie czynnika, mającego postulowane efekty z grupą osób niepoddanych temu działa-

⁸ Por. (Campbell i Stanley 1966).

niu, oraz badań typu oddziaływanie-kontrola, w których porównuje się wystarczająco liczną grupę wykazującą efekty postulowanego oddziaływania z grupą nieposiadającą tego atrybutu.

W badaniach przypadek-kontrola grupy dobierane są retrospektywnie, więc niemożliwa jest randomizacja. Zasadniczą techniką tu stosowaną jest dopasowanie, gdzie każdy przypadek (lub grupę) dopasowuje się do niepoddanej oddziaływaniami osobie, która ma taką samą wartość zmiennych kontrolnych (Susser 1973, 76-77; Rosenbaum 2002). Jako przykład takiego postępowania Susser podaje badania nad wpływem silnego zanieczyszczenia powietrza (zmienna niezależna) na raka płuc (zmienna zależna) i bronchit (zmienna zależna). Zjawiska, które mają także wpływ na zachorowalność na raka płuc i bronchit, to palenie papierosów oraz wiek (zmiennie kontrolne). Zmiennie kontrolne ustalano na stałym poziomie, analizując relacje między zmienną niezależną a zależnymi dla poszczególnych kategorii lub wartości zmiennych kontrolnych (np. częstość występowania raka płuc w każdym stopniu natężenia skażenia powietrza analizowano dla każdej kategorii palaczy, określonej przez ilość i rodzaj palenia). Innym sposobem jest standaryzacja rozkładu zmiennych kontrolnych w grupie badanej i kontrolnej (Susser 1973, 79-80).

W przypadku badań kohortowych, które są badaniami długofalowymi (np. ze względu na okres, który dzieli oddziaływanie czynnika od ujawnienia się jego skutków), łatwiej jest zapewnić, że obie grupy porównywane w badaniach są wzięte z tej samej populacji, jednak powstaje problem ustalenia wpływu na zmienną zależną z uwagi na upływ czasu między oddziaływaniem czynnika a ujawnieniem się skutków.

Podsumowując rozważania dotyczące sposobu organizacji badań i zbierania obserwacji, Susser podkreśla zalety planowania eksperymentalnego przede wszystkim pod względem kontroli czynników zewnętrznych w stosunku do rozważanej zależności przyczynowej. Jednak nad tymi zaletami w przypadku problemów badanych w epidemiologii przewagę mają studia obserwacyjne, gdyż

przewagę studiów eksperymentalnych pod względem kontroli czynników zewnętrznych oraz jakości miar wyrównuje przewaga studiów obserwacyjnych w kwestii reprezentatywności populacji i naturalności układu, a przede wszystkim w kwestii wykonalności poszukiwania determinantów. (Susser 1973, 88-89)

Druga strategia, po takim zaplanowaniu badań, by usunąć czynniki wpływające na badaną zależność przyczynową, dotyczy sposobu konstru-

owania modelu tej zależności. Ponieważ nie można poddać kontroli ani uwzględnić wszystkich zmiennych, konieczne więc jest wprowadzenie upraszczających założeń dotyczących tego, że nieuwzględnione zmienne nie mają wpływu na modelowaną zależność. Potrzebna w związku z tym jest strategia, która pozwoli ocenić słuszność takich założeń po skonstruowaniu modelu.

Ta strategia, zdaniem Sussera, zasadza się na wyróżnieniu rodzajów zmiennych niezależnych, które oddziałują na zmienną zależną. I tak, wskazuje on cztery podstawowe rodzaje takich zmiennych: hipotetycznie przyczynowa, kontrolowana, niekontrolowana oraz zmienna konfundująca.

Hipotetycznie przyczynowa zmienna jest przedmiotem badań zmierzających do wykrycia, w jakim stopniu zmiany w zmiennej zależnej można przypisać zmianom w tej badanej zmiennej niezależnej. Odpowiednie zaplanowanie badań oraz analiza mają pozwolić na wyeliminowanie wpływu innych czynników na zmienną zależną, a „precyzja tego efektu, który zostanie wyizolowany, będzie zależna częściowo od klarowności, z jaką niezależna badana zmienna i jej efekty są pojmowane, zdefiniowane i zmierzone” (Susser 1973, 92).

Zmienna kontrolna jest taką zmienną niezależną różną od hipotetycznie przyczynowej zmiennej, która spełnia dwa warunki: ma potencjalny wpływ na zmienną zależną oraz jest poddana kontroli w analizie. Kontrola może dokonywać się poprzez sposób zaplanowania badań (np. określenie zmienności w badaniu chorób dróg oddechowych, która jest spowodowana paleniem papierosów, przez badanie wpływu zanieczyszczenia powietrza na niepalaczy lub osiągnięcie jednakowego rozkładu palaczy w obu porównywanych grupach poprzez losowe przypisanie członków do grup lub dopasowanie). Może także dokonywać się poprzez analizę:

Źródła zewnętrznej zmienności w zmiennej zależnej są najpierw konceptualizowane w planowaniu badania jako zmienne kontrolne, tak aby mogły być określone, zmierzone i dzięki temu poddane analizie. (Susser 1973, 93)

Zmienne niekontrolowane także mogą wpływać na zmienną zależną, jednak nie są kontrolowane w analizie. Nie są też związane z hipotetyczną zmienną przyczynową, co różni je od konfunderów. Ogólnie więc można scharakteryzować te zmienne jako takie, które wyjaśniają zmienność zmiennej zależnej, która nie jest wyjaśniona w analizie, np. w przypadku raka płuc uwzględnienie takich zmiennych, jak palenie, zanieczyszczenie powietrza, wiek, płeć i klasa społeczna, pozostawia niewyjaśnione przypadki zachorowań osób, które pod żadnym z wymienionych względów nie należą do grupy

ryzyka i nie są wystawione na działanie tego rodzaju czynników. I na odwrót, nie wszystkie osoby wystawione na działanie wszystkich tych czynników chorują na raka płuc. Takie przypadki różnią się więc od pozostałych pod nieznanym względem.

Susser sugeruje, że jest kilka możliwych źródeł takiej niekontrolowanej wariancji:

Resztowa wariancja w zmiennej zależnej implikuje, że działają czynniki inne niż te związane ze zmienną badaną i zmiennymi kontrolowanymi. Ta wariancja może pochodzić od nieznanymi czynników w środowisku, nieprawidłowości (ang. vagaries) osobniczego rozwoju danego podmiotu (ang. host) oraz predyspozycji genetycznej lub czynników konstytucyjnych, które sprawiają, że ludzie stają się nadmiernie narażeni na daną chorobę. Dochodzi do tego także błąd pomiaru. (1973, 94)

Takie niekontrolowane źródła zmienności są opanowywane przez techniki statystyczne w planowaniu badań i analizie danych. Ponadto, jak podkreśla Susser, „W wielozmiennowych technikach analizy, jak wielokrotna regresja i analiza ścieżek, do równań dołącza się terminy, które pozwalają na niewyjaśnioną wariancję” (1973, 94).

Susser określa konfunder jako „zmienną niezależną, która powoduje zmianę w zmiennej zależnej i która zmienia się systematycznie z badaną hipotetyczną zmienną przyczynową” (1973, 95). Ze względu na ten systematyczny związek między hipotetyczną zmienną przyczynową a konfunderelem, może zaistnieć sytuacja taka, gdzie zmiany obserwowane w zmiennej zależnej dają się całkowicie wytłumaczyć przez zmiany w zmiennej konfundującej, a związek między hipotetyczną zmienną przyczynową a zmienną zależną jest pozorny. Przykładem, jaki mógłby to ilustrować, jest zdaniem Sussera możliwa – i założmy, że pozorna – zależność między zanieczyszczeniem powietrza (hipotetyczna zmienna przyczynowa) a bronchitem (zmienna zależna). Konfunderelem w tym przypadku może być gęstość zaludnienia, która z jednej strony ułatwia przenoszenie tej choroby, a z drugiej wpływa na zwiększenie zanieczyszczenia powietrza. Tym samym więc konfunder wyjaśnia wcześniej obserwowaną zależność między zanieczyszczeniem powietrza a bronchitem i ta zależność okazuje się pozorna.

Jak podkreśla Susser, konfundowanie może być ograniczone przez zarówno odpowiednie planowanie badań (dyskutowane w pierwszej strategii wykluczenie poza system lub zneutralizowanie przez taki sam rozkład w porównywanych grupach), jak i analizę danych.

W przypadku analizy zmienną konfundującą przekształca się na inny rodzaj zmiennej – kontrolowaną lub nie. Losowe przypisanie do jednej z dwóch grup w badaniu eksperymentalnym przekształca konfunder w zmienną niekontrolowaną, która tłumaczy wariancję resztową o takim samym rozkładzie dla każdej z grup. Dopasowanie w studium obserwacyjnym w podobny sposób przekształca tę zmienną w zmienną niekontrolowaną o takim samym rozkładzie w obu grupach.

Poprzez prowadzenie analizy dla różnych wartości zmiennej konfundującej, można zamienić ją w zmienną kontrolowaną. Jako przykład Susser podaje badania nad związkiem zachorowań na raka prostaty z zanieczyszczeniem powietrza, gdzie zmienną konfundującą jest status społeczno-ekonomiczny. Dla dwóch wyróżnionych grup (o wysokich i niskich dochodach) stwierdzona została ta sama asocjacja, co związane jest z zamienieniem konfundera w zmienną kontrolowaną.

Jak podkreśla Susser, może być sytuacja taka, gdy nawet losowe przypisanie do grupy w badaniu eksperymentalnym nie usunie konfundera. Przywołując sytuację zależności między zanieczyszczeniem powietrza a bronchitem, można byłoby sobie wyobrazić eksperyment polegający na losowym przypisaniu badanych osób do miejsca zamieszkania – o wysokim bądź niskim zanieczyszczeniu powietrza. Jednak takie losowe przypisanie nie zmieniłoby faktu, że w taki sam sposób zostałaby przypisana gęstość zaludnienia, co nie prowadziłoby do uniknięcia problemu konfundowania przez to ostatnie zjawisko.⁹

Bliźnięta jednojajowe mają dokładnie znany stopień konfundowania genetycznego (cechy genetyczne są dopasowane i całkowicie zneutralizowane)

⁹ Susser sygnalizuje także możliwość, zasugerowaną przez F. Galtona w 1875 roku, aby wykorzystać zjawisko konfundowania w planowaniu badań. Zasadnicza idea polega na wskazaniu czynników wrodzonych i środowiskowych poprzez badanie różnicy między bliźniętami jednojajowymi a dwujajowymi. Bliźnięta jednojajowe mają ten sam materiał genetyczny i pod tym względem wydają się doskonale dopasowane. Bliźnięta dwujajowe mają taki rozkład materiału genetycznego, jak inne rodzeństwa, ale znacznie bliższe uwarunkowania środowiskowe. Otrzymalibyśmy więc dwa równania: $BD - BJ = \text{efekt genetyczny}$, $Rod - BD = \text{efekt środowiskowy}$. Jednak, jak podkreśla Susser, atrakcyjność tych badań znacznie obniża fakt, że dwie trzecie jednojajowych bliźniąt dzielą tę samą kosmówkę i krążenie łożyskowe, a ze względu na zaburzenia wspólnego im dostarczania krwi mogą znacznie różnić się od bliźniąt dwujajowych pod względem doświadczeń prenatalnych. Po urodzeniu stopień zgodności doświadczeń bliźniąt jedno- i dwujajowych może się nawet odwrócić. Nie można więc przyjąć, że konfundowanie przez wspólne doświadczenia ma taki sam rozkład dla tych dwóch rodzajów bliźniąt, co zdaniem Sussera znacznie obniża atrakcyjność takich badań w studio efektów genetycznych (Susser 1973, 98-99).

i dlatego stanowią szczególnie cenny przedmiot obserwacji wpływu środowiska.

Susser zwraca uwagę na szczególnego rodzaju zmienną konfundującą, jaką jest zmienna zniekształcająca. Jest to konfunder, który jest skorelowany ze zmienną zależną w sposób przeciwny do hipotetycznej zmiennej przyczynowej. Stąd zmienna zniekształcająca może odwrócić, zmniejszyć, a nawet ukryć zależność między hipotetyczną zmienną przyczynową a zmienną zależną. Typowym przykładem zmiennej zniekształcającej jest wiek. Szczegółowiej omówiony przez Sussera przykład dotyczy pozornie pozytywnego wpływu stanu małżeńskiego na psychozę depresyjną. Później okazało się, że w rzeczywistości, gdyż w życiu małżeńskim często występują stany napięć. Jednak wzrost wieku jest pozytywnie skorelowany zarówno z byciem w stanie małżeńskim, jak i ryzykiem psychozy depresyjnej. Gdy dokonano analizy dla poszczególnych grup wiekowych, ta korelacja się odwróciła – częstość psychozy wśród niezamężnych była wyższa niż wśród zamężnych osób (Susser 1973, 101).¹⁰

Trzecia strategia jest kontynuacją poprzedniej, przy czym zamiast eliminować zmienne zewnętrzne lub nieistotne, analizuje się związek dwóch zmiennych powiązanych – przynajmniej hipotetycznie – związkiem przyczynowym w odniesieniu do innych zmiennych uwzględnianych w modelu. Biorąc pod uwagę zmianę własności uporządkowania czasowego oraz siły asocjacji, Susser wyróżnia pięć podstawowych funkcji, jakie mogą pełnić takie dodane zmienne: prosty poprzednik, zmienna wyjaśniająca, zakłócająca (ang. *intervening*), moderująca (ang. *moderator*) oraz składowa (ang. *component*). Jeśli hipotetyczna zależność ma postać $X \rightarrow Y$, to dwa pierwsze z wymienionych przypadków są zmiennymi na miejscu Z w schemacie: $Z \rightarrow X \rightarrow Y$. W trzecim zmienna leży pomiędzy zmienną zależną a niezależną: $X \rightarrow Z \rightarrow Y$. Zmienna moderująca działa w dowolnym momencie poprzedzającym w czasie zmienną zależną i w stosunku do zmiennej niezależnej może być ona poprzedzająca, równoczesna lub pośrednicząca. Natomiast zmienna składowa jest równoczesna ze zmienną niezależną, oddziałując na zmienną zależną.

¹⁰ Susser omawia jeszcze przypadek innego rodzaju konfundera, jakim jest zmienna dławiąca. Najczęściej chodzi o taki przypadek, gdy hipotetyczna zmienna przyczynowa jest zbyt złożona (wiele wymiarów lub kategorii) i tylko pewne jej elementy (np. jedno ze składowych zjawisk lub pewien przedział wartości) są związane relacją przyczynową ze zmienną zależną, a pozostałe są w ogóle nieistotne (Susser 1973, 104-110).

Dodanie prostego poprzednika (jako zmiennej kontrolowanej) nie wpływa na stwierdzoną pierwotnie zależność między zmienną niezależną a zależną, np. uwzględnienie diety matki nie zmienia zależności, jaka zachodzi między jej wagą a urodzeniową wagą dziecka. Nie wyjaśnia więc też zależności, jaka zachodzi między tymi zmiennymi, jedynie podaje wcześniejszy w czasie element „łańcucha przyczynowego”.

Jeśli pierwotnie stwierdzono zależność między X a Y , to dodanie zmiennej wyjaśniającej wykazuje, że ta zależność była pozorna: $X \leftarrow Z \rightarrow Y$ i jest wyjaśniona przez zmienną wyjaśniającą Z . Przykładem podanym przez Sussera jest wyjaśnienie zależności między ilością spożywanych kalorii przez matkę a wagą urodzeniową dziecka przez wagę matki (1973, 120).

Wprowadzenie zmiennej zakłócającej powoduje zniknięcie lub znaczące zredukowanie asocjacji między X a Y , które prowadzi do „interpretacji” tej asocjacji (Lazarsfeld 1961/1968). Susser podkreśla podobieństwo ról zmiennej zakłócającej i wyjaśniającej:

Efekt jest taki sam, jak w przypadku zmiennej wyjaśniającej. Ciąg przyczynowy tych zmiennych uformowany przez dodanie zmiennej zakłócającej do zaobserwowanej asocjacji między dwiema zmiennymi ma takie same własności, w zasadzie, jak ten uformowany przez dodanie zmiennej wyjaśniającej.

Do ich odróżnienia nie wystarczy sama analiza:

Zatem sama analiza nie może dać odpowiedzi na pytanie, czy trzecia zmienna jest zmienną, która „wyjaśnia” asocjacje, czy też jest zmienną zakłócającą, która „interpretuje” asocjacje. (Susser 1973, 122)

Potrzebne są dodatkowe informacje, np. o uprzedniości czasowej, lub dalsze badania. Susser nie artykułuje jasno różnicy między zmienną zakłócającą a prostym poprzednikiem, wydaje się jednak, że zasadnicze znaczenie ma to, czy badana hipotetyczna zmienna niezależna pozostaje zmienną niezależną po wprowadzeniu dodatkowej zmiennej (zmienna zakłócająca), czy też staje się zmienną zależną w stosunku do niej (prosty poprzednik).

Zmienna moderująca określa warunki, w jakich dotąd analizowana zależność między zmiennymi staje się mocniejsza lub słabsza: $X \rightarrow Y \leftarrow Z$. Jako przykład Susser podaje raka oskrzeli, który znacznie częściej występuje u palaczy narażonych na działanie rakotwórczego azbestu niż u pozostałych. Jak zaznacza, „*zmienna moderująca* może być zdefiniowana więc jako zmienna niezależna, która w połączeniu z dowolną inną zmienną niezależną wytwarza efekt wykraczający ponad ten, który każda z tych zmiennych wytwarzałyby z osobna” (Susser 1973, 126). Za P. Lazarsfeldem Susser określa wprowadzenie zmiennej moderującej do analiz jako specyfikację,

prowadzącą do wskazania różnych poziomów korelacji między badaną hipotetyczną zmienną przyczynową a zmienną zależną dla wszystkich wyszczególnionych wartości zmiennej moderującej.

Piąta kategoria, jaką stanowią zmienne składowe, opisuje elementy składające się na „globalną zmienną niezależną” (Susser 1973, 130). Taki rodzaj zmiennych niezależnych, jak zwraca uwagę Susser, często występuje w początkowej fazie badań, a dalsze kroki analizy – polegające właśnie na wyróżnieniu zmiennych składowych – mają określić, w jakiej mierze zaobserwowana asocjacja może być przypisana poszczególnym składowym. Przykładem takiej globalnej zmiennej niezależnej jest klasa społeczna, która może być definiowana, jak podkreśla Susser, przez połączenie wskaźników takich, jak zawód, wykształcenie i miejsce zamieszkania. Każdy z tych wskaźników jest więc zmienną składową.

Czwarte miejsce wśród wyróżnionych przez Sussera strategii zajmuje wykorzystanie pojęcia prawdopodobieństwa, a więc strategia związana przede wszystkim z analizą statystyczną dostępnych danych. Jako zasadniczy powód przywołania pojęcia prawdopodobieństwa, Susser podaje niewyjaśnioną w analizie resztową zmienność zmiennej zależnej.¹¹ Mimo to metoda statystyczna ma, zdaniem Sussera, kilka istotnych zastosowań w nieeksperymentalnych badaniach zależności przyczynowych. Umożliwia ona kontrolowanie w analizie niewyjaśnionej zmienności zmiennej zależnej przez estymację losowych prawdopodobieństw zdarzeń. Stopień zmienności przypisywany niekontrolowanym i nieznanym źródłom staje się podstawą obliczanych wartości oczekiwanych danej zmiennej. Następnie można określić, czy zaobserwowane wartości różnią się od oczekiwanych w granicach dopuszczalnej losowej zmienności i przetestować hipotezę statystyczną.

Metodę statystyczną można też wykorzystać, jak zaznacza Susser, do kontrolowania zmienności, która nie jest przypisywana losowi i daje się wyjaśnić. Wprowadza się w tym celu estymator „siły” planu badawczego, czyli prawdopodobną częstość, z jaką dane badanie – przy ustalonej wielkości efektu, wielkości próby i poziomie istotności – może ujawnić efekt, który nie jest losowy.

¹¹ Susser zaznacza, że są możliwe tu dwa podejścia: „pogląd deterministyczny przypisuje tę niewyjaśnioną zmienność czynnikom, które jeszcze należy wskazać”, natomiast możliwe jest też przypisanie jej przypadkowi (Susser 1973, 136). Te dwa podejścia określiłem wcześniej odpowiednio jako losowość epistemiczną oraz ontologiczną.

Z kolei techniki regresji wielokrotnej¹² pozwalają na estymację stopnia, w jakim określona zmienna niezależna jest odpowiedzialna za zmiany obserwowane w zmiennej zależnej. Tym samym ułatwiają ocenę zależności przyczynowej, określając siłę asocjacji między tymi zmiennymi.

Metodę statystyczną można także wykorzystać do skwantyfikowania i przetestowania modeli przyczynowych. W ten sposób można ocenić dopasowanie wartości przewidywanych w oparciu o dany model – przy różnych warunkach – z danymi zaobserwowanymi. W ten sposób można ocenić zależności między poszczególnymi zmiennymi w danym modelu, ale także dopasowanie do danych różnych alternatywnych modeli.

W podsumowaniu Susser przestrzega przed nieuważnym posługiwaniem się metodą statystyczną, która wprowadza szereg założeń do analizy, dotyczących m.in. postaci rozkładu, rzetelności pomiaru, kierunku zależności między zmiennymi itp. Jego zdaniem bardziej przekonujących rezultatów dostarcza długofalowa obserwacja tych samych zjawisk w tej samej populacji, zwłaszcza przy określaniu porządku czasowego (Susser 1973, 138).

Ostatnia z dyskutowanych przez Sussera strategii odwołuje się do „kryteriów osądu”, który ma dokonać „ważenia dostępnych dowodów empirycznych” (Susser 1973, 140-41). Do dyskusji wybrał on wersję poprzedzającą standardy podane przez Hilla, a mianowicie kryteria, które znalazły się w raporcie *Smoking and health* (1964). Są to już wcześniej przywoływane w niniejszej książce niektóre z kryteriów uwzględnionych przez Hilla: czasowa sekwencja zmiennych, spójność asocjacji przy powtórzeniach, jej siła, specyficzność i koherencja.

Dyskusja tych kryteriów przez Sussera w zasadniczych punktach jest zbieżna z analizami Hilla. Warto zwrócić jednak uwagę na kilka bardziej szczegółowych obserwacji. W odniesieniu do czasowej kolejności zjawisk Susser (1973, 142-48) zwraca uwagę na szereg trudności związanych z określeniem tego porządku, np. w kontekście zjawisk długotrwałych lub rozciągniętych w czasie (schizofrenia a status społeczno-ekonomiczny), zmieniających się z upływem czasu (choroba wieńcowa a osobowość) czy badanych retrospektywnie.

Wiedza o porządku czasowym jest także zależna od użytej miary częstości występowania zjawisk. W przypadku badań epidemiologicznych najczęściej stosowane to: zachorowalność, która wyraża częstość zachorowań w określonym okresie czasu, oraz prevalencja, która opisuje rozmiary zjawie-

¹² Por. (Brzeziński 343-430).

ska w danej populacji w określonym czasie, lecz niezależnie od momentu ich rozpoczęcia. Ponieważ zachorowalność daje możliwość ustalenia stanu innych zjawisk w momencie rozpoczęcia się danych przypadków badanej choroby, więc istnieje także możliwość ustalenia kolejności czasowej. Możliwości takiej nie daje jednak druga miara częstości, gdyż w przypadku prewalencji zmienna zależna i niezależna zaistniały przed momentem dokonania pomiaru (obserwacji).

W odniesieniu do spójności danej asocjacji w powtórzeniach Susser podkreśla, że w badaniach nieeksperymentalnych alternatywą dla eksperymentalnego dokładnego powtórzenia warunków jest spójność wyniku w powtarzanych obserwacjach. Siła tego kryterium zależy od zróżnicowania tych warunków, w których dokonuje się kolejnych obserwacji. Jako wzorcowe przykłady Susser podaje testowanie przez Snowa hipotezy o mechanizmie przenoszenia cholery w bardzo zróżnicowanych warunkach, dla których jedyną wspólną cechą okazał się ten właśnie mechanizm, także badania Hilla nad paleniem jako przyczyną raka płuc i E. Durkheima nad związkiem między ilością samobójstw a wyznaniem protestanckim (Susser 1973, 148-49).

Istotna uwaga, jaką czyni Susser, dotyczy tego, że spełnienie kryterium spójności nie gwarantuje, że w analizie zostaną usunięte konfundery. I tak, hipotetyczną przyczyną samobójstw mogły być nie przekonania religijne, lecz ściśle z nimi związana procedura pochówku – ponieważ odmawiano samobójcom pochówku zgodnego z rytuałem danego wyznania, więc możliwym powodem samobójstw mogła być niechęć do takiego właśnie rytuału (Susser 1973, 149). Gdyby tak było, spójność otrzymanych przez Durkheima wyników nie usunęłaby tego konfundera.

W odniesieniu do siły asocjacji Susser podkreśla, że w niektórych przypadkach, przy uwzględnieniu porządku czasowego, może być ona warunkiem wystarczającym stwierdzenia zależności przyczynowej między badanymi zjawiskami. I tak, w przypadku palenia papierosów, jak za Hillem podkreśla Susser, możliwy konfunder tej zależności musiałby występować ponad dziewięć razy częściej wśród palaczy niż niepalących i ponad dwadzieścia razy częściej wśród palących dużo, co z pewnością byłoby dającym się zauważyć zjawiskiem.

Susser dyskutuje także różne miary siły asocjacji. Jedną z nich jest tzw. ryzyko względne, które jest stosunkiem zachorowalności wśród tych, którzy są narażeni na działanie danego czynnika, do tych, którzy nie są. Ryzyko wskazuje natomiast, jaką bezwzględną zachorowalność na daną chorobę można przypisać czynnikowi przyczynowemu. Oblicza się ją, odejmując

zachorowalność wśród osób nienarażonych na dany czynnik od zachorowalności osób narażonych na jego działanie. Jednak, jak zaznacza Susser, ta miara zależy nie tylko od samej siły związku, lecz także od ilości członków danej populacji, którzy chorują na daną chorobę (którym można przypisać daną cechę).

Inną miarą może być obliczenie współczynnika korelacji lub współczynnika regresji. Współczynnik korelacji wskazuje stopień, w jakim zmienna niezależna zmienia się wraz ze zmienną zależną. Natomiast współczynnik regresji wyraża średnią oczekiwaną zmianę w wartości zmiennej zależnej przypadającą na zmianę zmiennej zależnej o jedną jednostkę. Z kolei wielokrotna analiza regresji wskazuje, jaka część zmienności zmiennej zależnej może zostać wyjaśniona przez daną zmienną lub zbiór zmiennych, gdy wartości innych stałych ustala się na stałym poziomie.

Przez specyficzność zależności Susser rozumie „precyzję, z jaką wystąpienie jednej zmiennej pozwala przewidzieć wystąpienie innej” (1973, 152). Najpełniejszą realizacją tego kryterium jest zależność jedno-jednoznaczna, gdzie zmienna niezależna jest zarówno warunkiem koniecznym, jak i wystarczającym skutku obserwowanego w zmiennej zależnej. O jego realizacji świadczy przejaw, który ma tylko jedną przyczynę.

Komentując zarzuty, jakie na podstawie braku specyficzności J. Berkson wysunął przeciwko tezie, że palenie powoduje raka płuc, Susser zaznacza, że „Specyficzność zwiększa wiarygodność (ang. plausibility) wnioskowania przyczynowego, ale brak specyficzności nie podważa go” (1973, 153).

W odniesieniu do wcześniej omówionych strategii Sussera warto zauważyć, że specyficzność można wyrazić czy zwiększyć, upraszczając warunki obserwacji, odpowiednio planując badania oraz dokonując w analizie rozkładu złożonych zmiennych niezależnych na prostsze czynniki o bardziej dookreślonym wpływie na zmienną zależną (np. skategoryzowanie grup palaczy według rodzaju palenia oraz ilości wypalanych papierosów).

Na zwiększenie specyficzności mają wpływ także dodane w analizie zmienne. Umożliwia się bowiem w ten sposób usunięcie konfunderów oraz nierelevantnych składników z dalszej analizy. Dotyczy to przede wszystkim analizy zmiennych zakłócających, moderujących i składowych. Jest to szczególnie istotne w przypadku asocjacji, które są słabe, choć spójne w różnych badaniach.

Ostatnim z poruszanych kryteriów jest koherencja. Związana jest ona z wyjaśnianiem w powiązaniu z innymi znanymi faktami dotyczącymi danej choroby oraz badanego czynnika przyczynowego. To kryterium opiera się na

istniejących dotąd wiedzy i teorii. Przykładu takiej zgodności, zdaniem Sussera, dostarczają badania nad rakiem płuc, które okazały się zgodne z histopatologią tej choroby.¹³

3.3 K. Rothmana i S. Greenlanda krytyka kryteriów przyczyny

Odniesienie kryterium proceduralnego, podanego w Punkcie 2.3, do kryteriów Hilla i ich modyfikacji (Cox, Wermuth, Susser) poprzedzę przywołaniem radykalnej krytyki tych kryteriów, którą w odniesieniu do zbioru kryteriów poszerzonego przez Hilla przeprowadzili K. Rothman i S. Greenland (1998).

Podkreślają oni, że siła związku między zmiennymi jest istotną wskazówką przy interpretacji przyczynowej tego związku, jednak nie jest jej warunkiem koniecznym. Jako przykład podają bierne palenie i raka płuc, gdzie istnieje słaba, lecz z pewnością przyczynowa zależność. Innym kontrprzykładem – pokazującym, że siła asocjacji nie jest też warunkiem wystarczającym dla interpretacji przyczynowej – są zależności mocne, lecz nieprzyczynowe, jak często ma to miejsce w przypadku konfunderów. Przykładem jest mocna zależność między zespołem Downa a oceną pourodzeniową, która jest konfundowana przez wiek matki. „Mocna asocjacja – jak podkreślają autorzy *Causation and causal inference* – służy jedynie wyeliminowaniu hipotez, że ta asocjacja jest w całości pochodną jednego słabego niezmiernego konfundera lub innego źródła umiarkowanego obciążenia” (1998, 25).

Spójność, przejawiająca się w powtarzalności danej asocjacji obserwowanej w różnych populacjach i w różnych okolicznościach, jak twierdzą Rothman i Greenland, także nie świadczy jednoznacznie o interpretacji przyczynowej. Jej brak bowiem nie przesądza o braku zależności przyczynowej w przypadku takich zjawisk, gdzie przyczyna wywołuje skutek tylko w bardzo szczególnych okolicznościach. Przykładami podanymi przez Rothmana i Greenlanda są zakażenie wirusem HIV podczas transfuzji krwi (występuje tylko, gdy obecny jest ten wirus) czy syndrom szoku toksycznego (ang. toxic shock syndrome) u kobiet stosujących tampony (rzadko przy towarzyszących nieznanym okolicznościach). Podsumowują krytykę tego kryterium,

¹³ W tym miejscu Susser podkreśla również, że liniowość zależności między zmienną zależną a niezależną jest dodatkowym potwierdzeniem koherencji, aczkolwiek nie jest to warunek konieczny zależności przyczynowej (Susser 1973, 155).

stwierdzając, że służy ono wyłącznie wyeliminowaniu hipotez przypisujących asocjację czynnikowi, który zmieniałby się w różnych badaniach.

Specyficzność, rozumiana jako możliwość określenia jednego skutku dla danej przyczyny, autorzy ci krytykują najostrzej, stwierdzając, że jest „całkowicie nietrafnym (ang. invalid)”, „bezużytecznym i mylącym kryterium”. Zasadniczy powód jest taki, że „Na podstawie samych tylko logicznych przesłanek nie można oczekiwać, by przyczyny danego skutku nie miały innych skutków. [...] Wystąpienie jednego skutku nie znosi możliwości, że istnieje jeszcze inny skutek” (1998, 25).

Zgadzając się, że przyczyna musi poprzedzać w czasie skutek, kwestionują jednoznaczność interpretacji opartej na kryterium uprzedniości czasowej. Jeśli bowiem zmienna zależna jest wcześniejsza niż niezależna w niektórych badaniach, to nie oznacza to, że nie ma między nimi zależności przyczynowej w innych niż badane okolicznościach. I na odwrót, jeśli zachodzi między nimi zgodna z oczekiwaniami relacja czasowa, to nie można stąd wyciągać wniosków o zależności przyczynowej między nimi. Jednak jeśli zmienna niezależna nie jest w odpowiednich okolicznościach uprzednia wobec zmiennej zależnej, to można wyciągnąć wniosek, że nie ma między nimi zależności przyczynowej. To jednak, jak podkreślają Rothman i Greenland, jest kryterium wynikające z samej definicji przyczynowości i dlatego uznają je za jedyny warunek konieczny tej zależności (1998, 28).

Zdaniem Rothmana i Greenlanda, warunkiem ani wystarczającym, ani koniecznym nie jest także gradient biologiczny, czyli liniowa zależność typu dawka-odpowiedź między zmiennymi. Niekiedy jest tak, że taka zależność wymaga przekroczenia pewnego progu w wartości zmiennej zależnej i wówczas otrzymujemy krzywą w postaci litery J. Przykładem podanym przez autorów *Causation and causal inference* jest zależność między spożyciem alkoholu a śmiertelnością, która jest wyższa wśród osób niepijących niż umiarkowanie pijących i gwałtownie wzrasta dla pijących dużo. Z kolei w przypadku konfunderów, jak we wspomnianym przykładzie zespołu Downa i oceny pourodzeniowej, można otrzymać liniową zależność mimo braku zależności przyczynowej, gdyż odzwierciedla ona liniową zależność między wiekiem matki a zespołem Downa. Tak więc, istnienie związku niemonotonicznego między zmienną zależną a niezależną prowadzi do odrzucenia jedynie tych hipotez przyczynowych, które są na tyle specyficzne, że pozwalają na sformułowanie przewidywania o istnieniu monotonicznej krzywej dawka-odpowiedź.

Kolejne z kryteriów Hilla, czyli prawdopodobieństwo zależności przyczynowej na gruncie posiadanej wiedzy, zwłaszcza biologicznej w przypadku epidemiologii, Rothman i Greenland traktują jako subiektywne i nie oparte ani na podstawie logicznej, ani na danych.¹⁴

W odniesieniu do kolejnego kryterium, jakim jest koherencja z wiedzą przyrodniczą i biologiczną o danej chorobie, odróżniają za Hillem dwa przypadki. Brak informacji o takiej zgodności nie świadczy przeciw danej zależności przyczynowej, natomiast gdy jest informacja o braku takiej zgodności, wówczas należałoby odrzucić daną hipotezę o zależności przyczynowej. Jednak, jak podkreślają Rothman i Greenland, „należy zawsze pamiętać o tym, że informacja o braku zgodności może być mylna lub błędnie zinterpretowana” (1998, 27).

Komentując ósme na liście Hilla kryterium dowodów empirycznych pochodzących z eksperymentu, podkreślają wieloznaczność w jego sformułowaniu. Z dwóch możliwych przypadków: eksperymenty na ludziach i laboratoryjne eksperymenty na zwierzętach, wybierają te pierwsze jako bardziej odpowiadające pierwotnym zamierzeniom Hilla, który – ich zdaniem – zamierzał uwzględnić dowody pochodzące z interwencji mających zapobiec działaniu danego czynnika. Informacje uzyskane jednak w takich eksperymentach są, według Rothmana i Greenlanda, „logicznie rzecz biorąc, nie kryterium, lecz testem hipotezy przyczynowej, testem po prostu niedostępnym w bardzo wielu okolicznościach” (1998, 27). Ponadto, nawet jeśli można przywołać informacje pochodzące z takiego eksperymentu, mają one ograniczone zastosowanie: „Chociaż testy eksperymentalne mogą być znacznie silniejsze niż inne testy, nie są tak decydujące, jak często się sądzi, ze względu na trudności w interpretacji” (Rothman i Greenland 1998, 27). Przykładem ilustrującym to jest eksperyment polegający na osuszaniu bagien, co miało być dowodem, że wydobywające się z nich gazy powodowały malarię. Osuszeniu bagien towarzyszył bowiem znaczący spadek zachorowań na tę chorobę. Jednak zawsze, zdaniem Rothmana i Greenlanda, możliwych jest wiele wytłumaczeń otrzymanego wyniku i tak we wspomnianym przypadku takim wytłumaczeniem, które okazało się słuszne, było to, że

¹⁴ Krótko odnoszą się do bayesianizmu, jako stanowiska, które uwzględnia dotychczasową wiedzę, przypisując danej hipotezie stopień prawdopodobieństwa z niej wynikający. Zdaniem jednak Rothmana i Greenlanda, nie gwarantuje to obiektywności kryterium prawdopodobieństwa (Rothman i Greenland 1998, 26), chociaż nie przywołują tu żadnych zarzutów stawianych stanowisku bayesianizmu.

malarię przenoszą komary, a osuszenie bagien znacznie zredukowało ich liczbę.

Ostatnie z kryteriów – analogia – sprowadza się zdaniem autorów *Causation and causal inference* do sprawności samych naukowców, którzy potrafią znaleźć podobieństwo niemal w każdej sytuacji. W najlepszym więc przypadku, jak stwierdzają Rothman i Greenland, analogia może być źródłem dopracowanych hipotez o badanym związku między zmiennymi. Brak analogii jednak świadczy, ich zdaniem, co najwyżej o braku wyobraźni lub doświadczenia u naukowca, nie jest jednak w żaden sposób związana z prawdziwością hipotezy.

Na zakończenie swojej krytyki Rothman i Greenland podkreślają, że Hill nie traktował swoich „standardów” jako kryteriów i z pewnością nie są one warunkami koniecznymi.

3.4 Jakościowe kryteria przyczyny a kryterium proceduralne

Dyskusję tego, na ile zarzuty stawiane przez tych autorów dotyczą zaproponowanego przeze mnie proceduralnego kryterium zależności przyczynowej, poprzedzę dyskusją związków, jakie zachodzą między zaproponowanym w Punkcie 2.3 kryterium proceduralnym a listą kryteriów czy „standardów” Hilla w wersji wyjściowej oraz w omówionych powyżej jej modyfikacjach.

Zasadnicza strategia zestawienia kryterium proceduralnego ze standardami Hilla polega na określeniu roli tych ostatnich jako *sprawdzianów* tego, czy kryterium niezmienniczości jest spełnione w określonej sytuacji. Dyskutowane przez Hilla i innych autorów ograniczenia tych standardów pokazują, że nie ma tu jednoznacznej relacji i niejednokrotnie nieznaną będzie pełna odpowiedź na pytanie, czy spełniony jest warunek niezmienniczości. Niemniej, mimo praktycznych trudności w jego sprawdzeniu, wprowadzenie tego warunku jest sprawą zasadniczą dla dyskusji nad modelowaniem przyczynowości. Ustanawia bowiem jego podstawy teoretyczne oraz pozwala porządkować znane z dotychczasowej literatury sposoby sprawdzenia, czy rzeczywiście mamy do czynienia z zależnością przyczynową między badanymi zmiennymi.

Przejdźmy zatem do dyskusji kolejnych standardów z listy Hilla. Rola, jaką Hill upatruje w sile związku, wynika przede wszystkim z faktu, że musiałyby istnieć zjawisko towarzyszące zaobserwowanej zależności, które musiałyby ją konfundować. Ze względu jednak na siłę związku jest – zda-

niem Hilla – bardzo mało prawdopodobne, że taki konfunder nie zostałby zauważony w badaniach. Tym samym więc dla Hilla siła związku między zmiennymi jest wskaźnikiem niezmienniczości tej zależności przy uwzględnieniu innych zjawisk czy też przy podjęciu interakcji badawczej z innymi zjawiskami. Nie jest to jednak warunek wystarczający tego, że zachowana jest żądana niezmienniczość z uwagi na to, iż nie jest wykluczone, że istnieje taka – dotąd nie podjęta – interakcja, w której konfunder dotychczas zawsze występujący łącznie z rzekomo przyczynową zmienną zostanie wyizolowany.

Jeśli spełniony jest warunek niezmienniczości, to przejawem tego jest dyskutowana przez Hilla spójność zależności, obserwowana w różnych populacjach, w różnych okolicznościach oraz dla różnych jednostek (poziomów) analizy. Z kolei brak spójności może świadczyć zarówno o tym, że dana zależność nie jest niezmiennicza, jak i o tym, że model tej zależności, który ukierunkowuje podejmowane z badaną rzeczywistością interakcje, nie jest kompletny. Wskazany przez Rothmana i Greenlanda przykład braku spójności w sytuacji, gdy zakażenia krwi wirusem HIV podczas transfuzji dokonuje się tylko w sytuacji obecności tego wirusa, jest przykładem tego drugiego przypadku, gdy model nie jest kompletny i uwzględnia jako przyczynę zakażenia wyłącznie transfuzję krwi. Porównując więc sytuację, gdy dochodzi do zakażenia krwi przy transfuzjach, gdzie stwierdza się obecność wirusa, z innymi przypadkami transfuzji, w których nie ma obecności tego wirusa, nie można wywnioskować niezmienniczości zależności, gdyż pytanie o niezmienniczość będzie dotyczyło bądź obu zależności łącznie (między zakażeniem a transfuzją i obecnością wirusa), bądź jednej z nich przy ustaleniu wartości drugiej na określonym poziomie. Przypadek opisywany przez Rothmana i Greenlanda nie spełnia żadnego z tych dwóch warunków i dlatego nie dotyczy związku między niezmienniczością a spójnością.

Niezmienniczość zależności przyczynowej jest warunkiem wystarczającym specyficzności tej zależności – kolejnego standardu Hilla – pod warunkiem, że taka właśnie jest ta zależność. W przypadku więc, gdy dana zależność faktycznie jest specyficzna, lecz nie jest jako taka określona w modelu zależności przyczynowych, który teoretycznie ukierunkowuje interakcje badawcze, można wyprowadzić wniosek, że taki model nigdy nie może posłużyć do wykazania niezmienniczości badanej zależności. Kryterium specyficzności odnosi się więc w pierwszym rzędzie do tego, jaki rodzaj modeli może być poddawany sprawdzaniu przy określaniu, czy spełniony jest warunek niezmienniczości dla jednej z zależności uwzględnionych w tym mo-

delu. Tutaj także znajdują zastosowanie – wcześniej omawiane – uwagi Sussera dotyczące dookreślenia i specyfikowania zmiennych – próby rozłożenia globalnych zmiennych niezależnych na prostsze i określenia dla nich związku, jaki zachodzi między nimi a zmienną zależną. Ma to służyć wyeliminowaniu z modelu tych zmiennych (lub składowych zmiennej niezależnej), które są nieistotne dla badanej zależności przyczynowej.

Bez wchodzenia w filozoficzną argumentację trudno jest określić dokładnie, dlaczego czasowość jest elementem definicyjnym przyczynowości. Niemniej zgodzimy się w tym miejscu z Rothmanem i Greenlandem. Nie jest to warunek wystarczający niezmienniczości zależności między badanymi zjawiskami, lecz co najwyżej warunek konieczny. Jeśli bowiem dla różnych interakcji zmienia się porządek czasowy rzekomo przyczynowej zmiennej oraz zmiennej od niej zależnej, wówczas wśród badanych zjawisk można wskazać takie, że wzięwszy je pod uwagę, zależność między badanymi zjawiskami znika.

Biorąc pod uwagę gradient biologiczny, czyli monotoniczność zależności między zmienną zależną a niezależną, można argumentację Hilla uzupełnić następująco, uwzględniając odniesienie do kryterium niezmienniczości. Dla wszystkich krzywych wyrażających niemonotoniczną zależność między badanymi zmiennymi (np. o kształcie L, J czy U) można znaleźć taki przedział wartości zmiennej niezależnej, dla którego zaobserwowana zależność się zmienia. Dla tego więc obszaru można postawić hipotezę o oddziaływaniu czynników, które nie zostały uwzględnione, a które wpływają na zmianę tej zależności. Nie jest zatem wykluczona możliwość, że uwzględnienie tych hipotetycznie oddziałujących czynników ostatecznie doprowadzi do stwierdzenia, że zaobserwowana zależność jest konfundowana. Taka natomiast możliwość nie powstaje w przypadku zależności monotonicznej, gdyż dla dowolnego przedziału zmiennej niezależnej kształt tej zależności jest dokładnie taki sam. Z punktu widzenia więc niezmienniczości można podać takie właśnie wyjaśnienie, dlaczego zależność monotoniczna jest pełniej swego rodzaju testem (czy przejawem) niezmienniczości niż jest nią zależność niemonotoniczna.

Kolejnym z warunków na liście Hilla jest wiarygodność, rozumiana jako prawdopodobieństwo na gruncie wiedzy biologicznej. Takie kryterium, jak zauważył sam Hill, jest względne w stosunku do zastanego stanu wiedzy biologicznej. Trudno jest, jak zauważają Rothman i Greenland, precyzyjnie odróżnić je od kolejnego kryterium, jakim jest koherencja z wiedzą ogólnoprzyrodniczą i biologiczną. Jedną z możliwych eksplikacji tej różnicy, jak

sugerują wspomniani autorzy, byłoby stanowisko bayesowskie, które wiarygodność proponuje wyartykułować jako pierwotne prawdopodobieństwo hipotezy. Nawet jeśli uznać, że obiektywistyczna wersja stanowiska bayesowskiego daje się obronić przed zarzutami ze strony personalistycznych (subiektywistycznych) bayesianistów (Horwich 1982; Mortimer 110-24; Howson i Urbach 1993, 413-19), to i tak brak byłoby podstaw, żeby twierdzić, że takie pierwotne prawdopodobieństwo hipotezy ma obiektywne podstawy.¹⁵ Jeśli więc kryterium wiarygodności nie sprowadza się do kryterium koherencji, to wydaje się, że różnicę między wiarygodnością a koherencją najtrafniej artykułują Cox i Wermuth, stwierdzając, że wiarygodność interpretacji zależności między zmiennymi w kategoriach przyczynowości zwiększa przede wszystkim podanie – zwykle dla innej jednostki analizy – mechanizmu, który łączy dwie badane zmienne. W badaniu cholery przez Snowa byłoby to podanie mechanizmu rozwoju choroby, który – na niższym poziomie analizy – łączy dwa badane zjawiska, jakimi są wprowadzenie do organizmu trucizny cholery oraz manifestacja jej objawów u chorego. Taki test niezmienniczości, polegający na wyeksplikowaniu mechanizmu łączącego dwie zmienne powiązane zależnością przyczynową, jest szczególnym przypadkiem testu spójności, gdyż nieuchronnie sprawdza się w ten sposób istnienie tej zależności dla innej populacji i w całkowicie innych okolicznościach.

Ósme z kryteriów, które podał Hill, to dowody empiryczne z interwencji. W każdej sytuacji, gdy takie zbliżone do eksperymentalnego badanie jest dopuszczalne, te dowody dostarczają najlepszego testu dla warunku niezmienniczości. W badaniach nieeksperymentalnych najczęściej jednak próba uzyskania takich dowodów z eksperymentu będzie prowadziła do znacznie gorszych sprawdzianów niezmienniczości niż studia obserwacyjne czy innego rodzaju badania. Wiąże się to przede wszystkim z faktem, że wprowadzenie badania eksperymentalnego w nieodpowiednim kontekście prowadzi do zredukowania liczby możliwych do podjęcia interakcji z daną rzeczywistością, a tym samym do zniweczenia wielu z wcześniej możliwych testów warunku niezmienniczości, jako warunku koniecznego i wystarczającego zależności przyczynowej między zmienną zależną a niezależną.

¹⁵ Zasadniczą częścią typowej obrony obiektywności stanowiska bayesowskiego jest odniesienie do twierzeń o zbieżności (Earman 1992; Hendricks 2001, 150-52). Jest wiele zastrzeżeń, co do osiągnięcia tej zbieżności (np. dla różnych naukowców, którzy przypiszą hipotezie skrajne wartości prawdopodobieństwa) lub tempa, w jakim taka zbieżność miałaby być osiągnięta.

Taki argument potwierdzają także omówione bardziej szczegółowo w tym punkcie uwagi Sussera dotyczące zalet i wad badań eksperymentalnych i obserwacyjnych. Oprócz tego, że modele stosowane do badań eksperymentalnych wymagają wprowadzenia szeregu założeń (m.in. postać rozkładu, dokładność pomiaru, ukierunkowanie zależności), studia obserwacyjne mają przewagę pod względem „reprezentatywności populacji i naturalności układu, a przede wszystkim w kwestii wykonalności poszukiwania determinantów” (Susser 1973, 88-89).

Ostatnie z kryteriów, jakim jest analogia, słusznie krytykują Rothman i Greenland, jako jedno z kryteriów zależności przyczynowej, które można wykorzystać do sprawdzenia warunku niezmienniczości. To „kryterium” odnosi się bowiem do sposobu formułowania hipotezy w taki sposób, aby możliwie najpełniej spełniała pozostałe kryteria, i z pewnością dlatego brak go na liście kryteriów zaproponowanych przez Coxa i Wermuth oraz Sussera.

Na zakończenie przywołajmy istotną modyfikację listy standardów Hilla, jaką proponują Cox i Wermuth. Otóż dwa pierwsze miejsca na zmodyfikowanej przez nich liście zajmuje perspektywna i retrospektywna zgodność z podanymi dla danej dziedziny badań wyjaśnieniami. Te dwa warunki, rozbudowujące kryterium koherencji Hilla i Sussera, z punktu widzenia warunku niezmienniczości nie stanowią dodatkowego testu tego warunku, gdyż nie wskazują, jak odnieść badaną zależność między zmiennymi do interakcji, które mogłyby wskazać zmienność tej zależności. Warunek zgodności teoretycznej natomiast stanowi integralną część pojęcia interakcji prowadzącej do ustalenia zależności przyczynowych między zmiennymi, w postaci jej *ukierunkowania teoretycznego*. To ukierunkowanie teoretyczne nie jest sprawdzianem, lecz nakłada ograniczenie na to, jakiego rodzaju interakcje są istotne z punktu widzenia sprawdzania, czy spełniony jest warunek niezmienniczości interakcji.

Z podobnych względów pierwsza z omawianych przez Sussera strategii wykrywania zależności przyczynowych, jaką jest uproszczenie warunków obserwacji, jest także składową proponowanego tu pojęcia interakcji, a mianowicie interakcji *zorientowanej* na daną dziedzinę badań. Zgodnie z tym, jak Susser charakteryzuje strategię upraszczania warunków obserwacji, otrzymany efekt ma doprowadzić badacza do wyeliminowania zmiennych nieistotnych i konfundujących, czyli – w przyjętych tu kategoriach – ma dla badanej dziedziny rzeczywistości ustalić, na jakie zjawiska zorientowana jest interakcja badawcza. Zadanie więc właściwego zorientowania interakcji

badawczej nie jest kwestią sprawdzania warunku niezmienniczości, lecz odpowiedniego skonstruowania planu badawczego.

Na zakończenie tego punktu krótko odnieśmy się do krytyki kryteriów Hilla oraz ich modyfikacji. Rothman i Greenland zasadniczo przyjmują tradycyjne empirycystyczne założenie o dychotomiczności podziału informacji uzyskanych na drodze dedukcji oraz z doświadczenia. Zgodnie z przyjętą przez nich tradycyjną wersją metodologii empirycystycznej nie ma podstaw do utrzymywania innego niż hipotetyczne istnienia zależności przyczynowych między zmiennymi. W zasadzie więc ich stanowisko, zgodnie z dyskusją przeprowadzoną w Zakończeniu, waha się między realizmem, postulującym hipotetyczne istnienie procesu generującego, a antyrealizmem negującym możliwość stwierdzenia zachodzenia zależności przyczynowych między zmiennymi. Zdaniem Rothmana i Greenlanda kryteria zaproponowane przez Hilla, a odwołujące się do związków z doświadczeniem wykraczających poza wnioskowanie dedukcyjne, można traktować jedynie jako pomoc przy testach *negatywnych* hipotez, postulujących zależności lub procesy przyczynowe. Są to testy negatywne, gdyż można je wykorzystać tylko przy próbach obalania takich hipotez, natomiast są bezużyteczne przy próbach ich potwierdzenia.¹⁶ Jednak przy pytaniu o bliższą charakterystykę różnicy między postulowaną hipotetycznie zależnością przyczynową między zmiennymi a jedynie zaobserwowaną między nimi korelacją, Rothman i Greenland zmuszeni będą do dokładniejszego określenia swojego stanowiska, bądź w kierunku realizmu naukowego, bądź antyrealizmu. Z uwag przez nich poczynionych w tekście można wywnioskować, że nie postulują radykalnego odrzucenia pojęcia przyczynowości w badaniach nieeksperymentalnych ani też jego redukcji do pojęć obserwacyjnych. Zatem ostatecznie różnicę należałoby scharakteryzować podobnie, jak zaproponował to Goldthorpe, a mianowicie hipotezy o zależnościach przyczynowych między zmiennymi należy rozumieć w perspektywie hipotetycznych procesów generujących obserwowane wartości tych zmiennych.

W podstawowych założeniach krytyki Hilla kryteriów zależności przyczynowych między zmiennymi Rothman i Greenland pomijają rozwijaną w tej książce możliwość, jaką jest przyjęcie pierwotnego i nieredukowalnego charakteru pojęcia interakcji. W związku z tym ich krytyczna argumentacja, opierająca się na dychotomiczności logiki i doświadczenia, traci swoją zasadniczą podstawę, jaką jest przyjęcie hipotetycznego istnienia ge-

¹⁶ Podobne stanowisko zajmują G. A. Ferguson i Y. Takane; por. (1989/2003, 152, 425).

nerujących procesów przyczynowych. Wprowadzenie pojęcia przyczynowości w kategoriach pierwotnej i nieredukowalnej interakcji pozwala bowiem na wykorzystanie – w omówiony wcześniej sposób – kryteriów Hilla, w ich pierwotnej oraz zmodyfikowanej postaci do przeprowadzania testów tego, czy faktycznie spełniony w danej sytuacji jest warunek niezmienniczości badanej zależności z uwagi na możliwe interakcje z daną dziedziną rzeczywistości. Ponieważ jest to warunek konieczny i wystarczający ustalenia zależności przyczynowej między badanymi zmiennymi, można więc także twierdzić, że sprawdzenie poszczególnych testów przybliży nas do stwierdzenia, czy dana zależność faktycznie jest przyczynowa.

Zanim przejdziemy do bliższej analizy różnych współcześnie stosowanych sposobów modelowania zależności przyczynowych, podkreślmy dalszą konsekwencję tego, co powiedziano w poprzednim akapicie. Zgadzam się z wnioskiem Hilla oraz wielu innych autorów, że nie istnieje formalny test zależności między zmiennymi (a więc taki test, który nie odwoływałby się do własności podejmowanych z reprezentowanymi przez te zmienne interakcji), który byłby wystarczającym warunkiem ustalenia zależności przyczynowej między nimi. Rozwinięcie tej konsekwencji omówiłem, referując za Susserem rolę metody statystycznej w badaniach nieeksperymentalnych przy ustalaniu zależności przyczynowych. Jest wiele istotnych problemów, które można dzięki narzędziom formalnej analizy rozwiązać, zmierzając do ustalenia tych zależności, jednak żadne z nich nie dostarcza testu, który byłby wystarczający do ostatecznego ustalenia, że taka zależność faktycznie zachodzi.

4

Typy ilościowych modeli przyczynowych

Część pierwszą książki, poświęconą proceduralnemu kryterium zależności przyczynowych, zamyka dyskusja odnosząca je do kryteriów stosowanych współcześnie w matematycznych modelach przyczynowych. Rozważania te prowadzę w świetle pytania: „Czy można w pełni zautomatyzować odkrywanie zależności przyczynowych?”.

Takie pytanie wydaje się nie tylko istotne ze względów praktycznych, związanych z prowadzonymi badaniami nad zależnościami między poszczególnymi zjawiskami. Takie pytanie, jak wyjaśnię szczegółowo poniżej, posiada ważkie persupozycje filozoficzne, uzasadniane na gruncie stanowiska reliabilizmu.¹ W ten właśnie sposób pozytywną odpowiedź na sformułowane tu pytanie systematycznie rozwijają C. Glymour i jego współpracownicy, zwłaszcza P. Spirtes i R. Scheines. Efektem ich prac jest program TETRAD, którego zadaniem jest odkrywanie zależności przyczynowych bezpośrednio z danych, a przynajmniej określenie, jakie wnioski o zależnościach przyczynowych można wyprowadzić po wyraźnym określeniu przyjmowanych założeń i rodzaju dostępnych danych.

Z tego też względu prace zespołu kierowanego przez Glymoura są poniekąd centralnym tematem dyskusji w tym rozdziale. W punkcie pierwszym zaanonsowane zostaną najważniejsze ze stosowanych współcześnie sposobów modelowania zależności przyczynowych. Spośród nich najwięcej uwagi poświęcę modelom równań strukturalnych SEM (skrót ang. structural equation models) oraz modelom przyczynowym Rubina RCM (skrót ang.

¹ Genezę oraz podstawowe założenia tego stanowiska referuję w (2005); por. też (Juhl 1993; Jain i in. 1999).

Rubin causal models).² Zwróćę przy tym uwagę przede wszystkim na ich związek z graficznymi modelami przyczynowości, tworzonymi za pomocą sieci bayesowskich. Za zaletę tych ostatnich, w stosunku do pozostałych sposobów modelowania zależności przyczynowych, ich twórcy uznają możliwość automatyzacji procesu uaktualniania rozkładów prawdopodobieństwa w oparciu o dane napływające z doświadczenia.

4.1 Podziały modeli przyczynowych

W poniżej zaproponowanym przeglądzie sposobów modelowania przyczynowości skrzyżowane zostaną ich dwa podziały.³ Zasadniczy podział dotyczy tych sposobów modelowania, które wykorzystują kryteria wykraczające poza związki statystyczne (właściwe modele przyczynowe), oraz modeli statystycznych (statystyczne modele przyczynowe).⁴ Więcej miejsca poświęcimy tym pierwszym, te drugie charakteryzując tylko ogólnie.

Drugi wykorzystywany tu podział modeli zależności przyczynowych odnosi się do rodzaju wykorzystanych technik: graficznych lub niegraficznych. Reprezentacja graficzna zależności przyczynowych nie jest warunkiem koniecznym podania modelu przyczynowego, jednak znacząco ułatwia określenie jego własności oraz uogólnienie ich na klasę modeli, zwłaszcza równoważnych obserwacyjnie (jest to tzw. wzorzec).

Rozpocznę od krótkiego omówienia różnicy między modelami statystycznymi a przyczynowymi w sensie właściwym. Dokonując przeglądu najważniejszych metodologicznych podejść do modelowania zależności przyczynowych, Cox i Wermuth (2001) wyróżnili trzy podstawowe ich typy. Wyróżnikiem pierwszego podejścia jest sprowadzanie zależności przyczy-

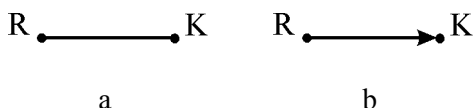
² Taką terminologię na określenie omawianego tu modelu rozpowszechnia szczególnie P. Holland, zwłaszcza w szeroko dyskutowanej i często przywoływanej pracy (Holland 1986b). Pearl proponuje inne określenie tego rodzaju modeli jako „modeli potencjalnych rezultatów” (ang. potential outcome models), przejmując w ten sposób terminologię J. Neymana; por. (Pearl 2000, 243-44).

³ Por. m.in. (Abbott 1998; Western 1999; Arjas 2001; Ash i in. 2001; Snijders i Hagenaars 2001).

⁴ Proponuję takie nazewnictwo dla uniknięcia pomieszania terminologicznego, jakie panuje w literaturze anglojęzycznej, gdzie mowa jest o statystycznych modelach przyczynowości, które następnie krytykuje się jako nieprzyczynowe, bądź mowa jest o statystycznych modelach, które krytykuje się za to, że nie są przyczynowe.

nowych do pewnego rodzaju zależności statystycznych⁵, a mianowicie takich, które są stabilne – lub odporne (Goldthorpe 2001, 1; Pan i Frank 2003; Robinson 2005), czyli w dużej mierze niezależne od konkretnego układu wartości parametrów modelu, jakie stwierdzono w konkretnym badaniu.

Założmy, że na podstawie danych jest jasne, iż wysokim wynikiem testów sprawdzających – dla określonego wieku badanych – umiejętności rachunkowe (R) towarzyszą wysokie wyniki testów kompetencji językowej (K). Jak za pomocą korelacji wyrazić wniosek, że R jest przyczyną K? Cox i Wermuth podkreślają (2001, 65), że korelacja – w przeciwieństwie do przyczynowości – jest zależnością symetryczną (por. Rysunek 4.1a), a więc jeśli R jest w tej relacji do K, to K jest w tej samej relacji do R. Jednak związek przyczynowy między R a K jest asymetryczny (por. Rysunek 4.1b) – jeśli jedno jest przyczyną drugiego, to nie na odwrót.



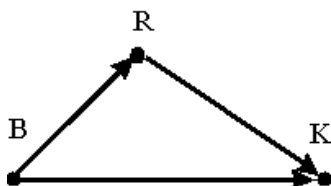
Rysunek 4.1. (a) symetryczna asocjacja R i K; (b) asymetryczna zależność przyczynowa R i K.

Cox i Wermuth rozważają dwa sposoby wprowadzenia asymetrii w korelacji R i K (2001, 65-66). Pierwszy polega na przyjęciu założenia, że zdarzenia występujące wcześniej (parametr czasu) pełnią funkcję eksplanacyjną w stosunku do zdarzeń późniejszych (np. pewne aspekty wykształcenia mogą być eksplanacyjnymi zmiennymi w odniesieniu do późniejszego przebiegu kariery zawodowej). W tym przypadku nie chodzi o czas, kiedy dokonano obserwacji, lecz o czas wystąpienia danych zdarzeń lub czas, do którego odnoszą się cechy będące przedmiotem badania.

Drugi sposób polega na wprowadzeniu roboczej hipotezy, opartej na zaakceptowanej już teorii lub danych empirycznych pochodzących z innych badań. I tak, jako ilustrację można przytoczyć dochód, który jest asymetrycznie związany z postawami względem różnych kwestii społecznych i politycznych, co oczywiście ma miejsce w kontekście bardziej złożonych relacji.

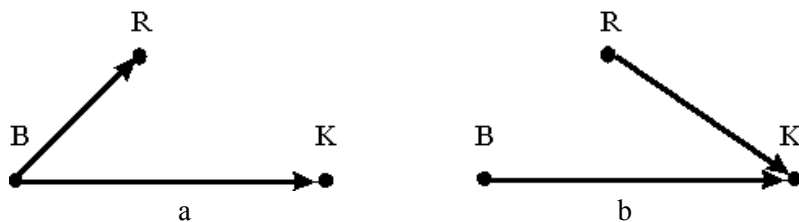
⁵ Mało przekonujące, ale możliwe byłoby stanowisko broniące tej redukcji tylko na płaszczyźnie metodologicznej, niezależnie od aspektu ontologicznego. Uzasadnieniem często tu przywoływanym jest ujawnianie się zależności przyczynowych w postaci korelacji.

Sytuacja komplikuje się, gdy uwzględnić większą liczbę zmiennych. Założmy więc, że mamy do czynienia z kolejną zmienną B, która może mieć wpływ przyczynowy na R oraz na K. Najliczniejszy zestaw zależności przyczynowych otrzymalibyśmy, gdyby B było przyczyną R oraz K, a także – niezależnie od B – R byłoby przyczyną K, jak ilustruje to Rysunek 4.2.



Rysunek 4.2.

Na Rysunku 4.2 B jest *przyczyną wspólną* R a K, natomiast K jest *skutkiem wspólnym* B i R. Te przypadki mogą wystąpić oddzielnie, co ilustruje Rysunek 4.3.



Rysunek 4.3. (a) skorelowane R i K, są nieskorelowane warunkowo względem B; (b) B i R nieskorelowane, są skorelowane warunkowo względem B.

Przypadek (a) na Rysunku 4.3, gdy między R i K brak krawędzi oznaczającej zależność przyczynową ilustruje typową sytuację, gdy do oszacowania przedstawionych zależności stosuje się analizę z wykorzystaniem pewnej postaci regresji (Cox i Wermuth 2001, 66).⁶ Jeśli nie uwzględni się zmiennej B, to w analizie otrzymamy zależność statystyczną między R a K, która jednak nie oznacza zależności przyczynowej, a może prowadzić do błędu konfundowania.

⁶ Bogaty przegląd historyczny podał R. Howarth (2001). Na ograniczenia metod regresji wskazują m.in. (Guilford 1954/2005; Clogg i Haritou 1997; Glymour 2001; Mynarski 2003).

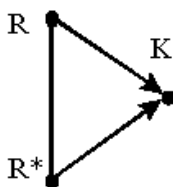
Autorami, którzy w szczególnie wyrazisty sposób wykorzystali powyższe wnioski w podanej przez siebie definicji zależności przyczynowej, są I. J. Good, P. Suppes, a w odniesieniu do szeregów czasowych – N. Wiener oraz C. W. Granger. Zgodnie z ich zamysłem R uznać można za przyczynę K, gdy zachodzi między nimi zależność statystyczna, a jej kierunek nie zmienia się przy uwzględnieniu innych zmiennych, które mogą mieć wpływ na R.

Cox i Wermuth podkreślają jednak, że taka definicja nie może być spełniona, bo w badaniach można co najwyżej odnieść ją do *zmiierzonych* zmiennych i sprawdzić, czy żadna z nich nie jest przyczyną wspólną R i K. Jak zaznaczają jednak:

Możliwość, że tę zależność dałoby się wyjaśnić przez zmienne eksplanacyjne w stosunku do [R], których nie zmierzono, a więc przez tzw. niezaobserwowane konfundery, jest tym mniej prawdopodobna, im silniejszy jest dostrzegalny efekt i może być uwzględniona tylko ze względu na ogólne argumenty odwołujące się do danej dziedziny badań. (2001, 66)

Pomocna w określaniu roli zjawisk, które nie są zmierzone, może być zwłaszcza statystyczna analiza wrażliwości. Określa ona, jakie własności musiałaby mieć niezaobserwowana zmienna będąca przyczyną wspólną R i K (konfunder), aby całkowicie wyjaśnić zależność między R a K. Kiedy odkryta w badaniach zależność empiryczna jest bardzo mocna, wówczas w zasadzie można uznać za nieprawdopodobne to, że istnieje niezmierna zmienna, która miałaby na tyle mocny wpływ na R lub K, aby podważyć zaobserwowany efekt (Rosenbaum 2002).

Jedną z trudności statystycznego podejścia do modelowania związków przyczynowych powstaje, gdy mamy do czynienia ze skorelowanymi, np. R oraz R*, mającymi wpływ przyczynowy na K, i badamy status przyczynowy R względem K.



Rysunek 4.4.

Jeśli zależność statystyczna między R a K pozostaje taka sama niezależnie od tego, czy warunkujemy ze względu na R*, czy nie, a więc niezależnie

od tego, czy R^* jest uwzględniane w równaniu regresji, nie powstaje zasadnicza trudność, przynajmniej jeśli chodzi o poziom kwalitatywny. Jak podkreślają jednak Cox i Wermuth, dość często spotyka się wyraźną zależność od obu zmiennych R i R^* łącznie, chociaż każda ze zmiennych wzięta osobno jest wystarczająca do wyjaśnienia tej zależności. W takim przypadku pozostają trzy możliwości. Po pierwsze, można potraktować R i R^* łącznie jako złożoną przyczynę K . Po drugie, można uznać istnienie dwóch przyczyn R oraz R^* . Po trzecie, można pozyskać dodatke informacje, aby dokładniej określić relację między zmiennymi R a R^* , np. R^* jako przyczynę R .

Przykładem mogą być wyniki testów edukacyjnych, mierzących sprawność rachunkową (R^*) i językową (R) dziecka w określonym wieku i określających ich efekt, czyli pewną cechę osoby dorosłej (K) (Cox i Wermuth 2001, 67). W tym przykładzie R^* i R nie pozostają w zależności przyczynowej, więc pozostają tylko dwie z wyżej wymienionych możliwości. Można potraktować wyniki testów jako zmienną dwuwymiarową i pominąć – przynajmniej tymczasowo – ewentualne różnice we wpływie sprawności rachunkowych i językowych na dorosłe życie. Można też przyjąć dwie interpretacje jako zgodne z danymi⁷ – jedną, odwołującą się do sprawności rachunkowej, a drugą – do kompetencji językowej. Każda opierałaby się na uważnym przestudiowaniu empirycznych równań regresji, np. dla binarnej zmiennej R przy użyciu regresji logistycznej.

Znaną i szeroko stosowaną w ekonometrii wersją podejścia statystycznego w modelowaniu zależności przyczynowych jest koncepcja noblisty C. Grangera (1969). Zależności przyczynowe wykrywane są – co podkreśla Goldthorpe (2001, 2) – na podstawie danych ekonomicznych, mających postać szeregów czasowych. Charakterystyczne dla tego podejścia jest wyeksponowanie mocy predyktywnej wskazanych w analizie zależności przyczynowych. Ogólnie kryterium Grangera można przybliżyć następująco: zmienna R jest *przyczyną w sensie Grangera*⁸ zmiennej K , jeśli mimo uwzględnienia wszystkich informacji z wyjątkiem wartości R , te wartości są pomocne przy przewidywaniu przyszłych wartości K .⁹

⁷ Cox i Wermuth stwierdzają: „uznalibyśmy dwa alternatywne proste wyjaśnienia, które są spójne z tymi danymi” (2001, 68).

⁸ W literaturze anglojęzycznej występuje termin techniczny *Granger causality* na oznaczenie zależności wskazanych na podstawie kryterium Grangera.

⁹ Bardziej szczegółowo dyskutują to kryterium m.in. B. Zacny (2002), E. Gatnar (2002, 95-103) oraz Z. Zieliński i J. Bruzda (2004).

Faktycznie możliwe jest jednak uwzględnienie tylko zaobserwowanych wartości zmiennych,¹⁰ a przyczyną w sensie Grangera jest każda niezerowa korelacja cząstkowa¹¹, która zwiększa możliwości predykcji. Jak argumentował P. Holland (1986a; 1986b), kryterium zależności przyczynowych w sensie Grangera pozwala przede wszystkim na wskazanie i wyeliminowanie pozornych zależności przyczynowych, które Granger określał jako „nieprzyczynowe”.¹²

Do znanych, zwłaszcza w socjologii, statystycznych sposobów modelowania zależności przyczynowych należy procedura *elaboracji* zaproponowana przez P. Lazarsfelda. Można ją scharakteryzować podobnie jak przyczynę w sensie Grangera. Lazarsfeld – w odróżnieniu od Grangera – był jednak zainteresowany określaniem pozornych zależności przyczynowych przede wszystkim ze względu na wyjaśnianie, a nie predykcję. W *elaboracji* badania rozpoczyna się od stwierdzenia korelacji między dwiema zmiennymi, która jest teoretycznie interesująca, np. między miejscem zamieszkania a sposobem głosowania. Następny krok polega na poszukiwaniu wyjaśnienia tej właśnie korelacji, zamiast zakładania, że jest to zależność przyczynowa. Wprowadza się zatem dodatkowe zmienne, np. klasa społeczna czy narodowość, które mogą w rezultacie dać zerową korelację cząstkową między zmiennymi, w przypadku których – jak R i K – domniema się zachodzenia zależności przyczynowej. Uznanie, że taka dodatkowa zmienna jest jednocześnie sama przyczyną zarówno R, jak i K, jest tymczasowe i podlega tej samej procedurze, jaka została zastosowana dla R i K.

Statystyczne podejście do modelowania przyczynowości Goldthorpe określa „pragmatycznym utylitaryzmem” (2001, 3) i ocenia negatywnie:

jeśli patrzy się na przyczynowość w ten właśnie sposób, to ustalenie, że zachodzi zależność przyczynowa, jest *wyłącznie* kwestią wnioskowania statystycznego, bez udziału jakichkolwiek szerszych rozważań. Przyczynowość można wyprowadzić bezpośrednio z analizy regularności empirycznych, stosując zasady, które aplikują się w taki sam sposób do różnych dziedzin badań i bez wymogu przywołania wiedzy o danym przedmiocie – czy to wiedzy ogólnej, czy ważniejszej – teorii. (Goldthorpe 2001, 3)

¹⁰ Por. (Siciński 1968).

¹¹ Pojęcie korelacji cząstkowej oraz metody jej analizy dla dowolnej liczby zmiennej podają m.in.: T. W. Anderson (2003), M. Dobosz (2001, 81-83). Zastosowanie w tworzeniu uproszczonych modeli przyczynowych prezentują A. de la Fuente i in. (2004).

¹² Ogólnie sens „nieprzyczynowej” zależności można określić następująco: R nie jest przyczyną w sensie Grangera zmiennej K ze względu na informacje W, o ile korelacja między R a K jest zerowa warunkowo ze względu na W; por. (Goldthorpe 2001, 2).

Nawet w odniesieniu do kryterium przyczyny w sensie Grangera, gdzie zasadniczą kwestią jest – niezwykle praktycznie użyteczna – predykcja, podnosi się tę trudność, że nie prowadzi ono do wyjaśnień przyczynowych.¹³ Aby podać takie wyjaśnienia, należałoby uwzględnić przede wszystkim wpływ ludzkich interwencji na zmiany w badanych zjawiskach, które „naturalnie nie zmieniają się samodzielnie, bez ludzkiej woli” (Sen 1986, 14).

Goldthorpe jest zdania, że jest to możliwe wyłącznie, gdy „wykazany związek statystycznym daje się podstawę teoretyczną, która ostatecznie sięgać będzie poziomu mikroekonomii” (Goldthorpe 2001, 3). Podobny wniosek sformułował A. K. Sen:

gdy „bezmisłna makroekonomia” może służyć jako podstawa przewidywań – a przynajmniej „prostych i bezpośrednich” – to każde „głębsze wyjaśnienie” zmian odnośnych wielkości może być ostatecznie uzyskane jedynie poprzez teorię i to takiego rodzaju, który odwołuje się do „celów, wiedzy, rozumowania i decyzji” jednostek działających w społeczeństwie. (1986, 14)

Zdaniem Goldthorpe’a, rola rozważań teoretycznych jest większa jeszcze w przypadku takich nauk, w których predykcje są niemożliwe lub rzadkie. Ilustruje to przykładem. Załóżmy, że w oparciu o jedno ze statystycznych kryteriów stwierdza się, że zachodzi zależność przyczynowa między posiadanym wykształceniem a poziomem dochodów. Traktowanie wykształcenia jako przyczyny wysokości dochodów byłoby, zdaniem Goldthorpe’a, niedopuszczalnym uproszczeniem. Ludzie zdobywają określone stanowisko, ponieważ inne osoby lub instytucje im je zaoferowały lub też sami je stworzyli, będąc uczestnikami rynku dóbr i usług. Jako uogólnienie Goldthorpe formułuje następujący zarzut:

Zwłaszcza z punktu widzenia indywidualizmu metodologicznego socjologowie poddali ostrej krytyce założenie, że techniki statystyczne mogą same dostarczać adekwatnych wyjaśnień przyczynowych zjawisk społecznych. Takie techniki mogą pokazać jedynie zależności między zmiennymi, a nie to, jak te relacje rzeczywiście zostały wytworzone – a w rzeczy samej mogą być wytwarzane tylko poprzez działanie i interakcje indywidualów. (2001, 3)

Większość metodologów przyczynowości zgodzi się z pierwszym orzeczeniem, drugie natomiast pozostaje kontrowersyjne i w zasadzie bez ja-

¹³ Wśród ekonomistów, a nawet ekonometrów, którzy podnoszą ten problem, są m.in. R. L. Basmann (1988) i A. Zellner (1988).

snego związku logicznego z pierwszym.¹⁴ Uzasadnieniem tego stwierdzenia ma być dalsza analiza podanego przykładu: „pozostaje pytanie, jak powstaje ta [odporna] zależność [wysokości dochodu od wykształcenia]?” (Goldthorpe 2001, 4). Można wziąć pod uwagę kilka możliwości: wykształcenie dostarcza wiedzy i umiejętności, które mają wartość rynkową, wykształcenie jest dla pracodawcy jedynie wskaźnikiem psychologicznej lub socjologicznej charakterystyki osoby poszukującej pracy, wykształcenie kształtuje podaż pracy w przypadku niektórych rodzajów zatrudnienia, itp. Ustalenie związku przyczynowego wymagałoby więc usytuowania tego związku w obrębie pewnego uogólnionego dyskursu dotyczącego działania, który byłby też konsekwencją dobrze rozwiniętej teorii działania społecznego.

Pearl, charakteryzując modele statystyczne przyczynowości, wskazuje na inny ich niedostatek. Zasadniczą cechą statystycznych modeli przyczynowości – w odróżnieniu od modeli przyczynowych w sensie właściwym – jest to, że uwzględniają one wyłącznie zależności statystyczne między zmiennymi (zwłaszcza warunkowe niezależności), a nie tkwiące głębiej zależności przyczynowe. W rezultacie więc nie można w oparciu o takie modele określić, czy dana zależność nie zmieni się pod wpływem określonej interwencji. Tym samym więc statystyczne modele nie pozwalają przewidywać skutków interwencji:

Czasowa uprzedniość wśród zmiennych może dostarczyć informacji o (braku) przyczynowych zależności – późniejsze zdarzenie nie może być przyczyną wcześniejszego. Indeksowane czasowo rozkłady [...] mogą być zatem uznane za przypadek graniczny między modelami statystycznymi a przyczynowymi. My jednak zaklasyfikujemy te modele jako statystyczne, ponieważ większość pytań dotyczących działań (ang. policy-related) *nie może* mieć odpowiedzi na podstawie takich rozkładów [...]. (Pearl 2000, 39)

Przechodząc do drugiego podejścia – modelowania zależności przyczynowych w sensie właściwym – należy podkreślić, że zasadnicza różnica, jaka dzieli je od podejścia statystycznego, polega na uwzględnianiu zależności kontrfaktycznych oraz interwencji. Goldthorpe podkreśla, że wzorcem dla tego podejścia są badania eksperymentalne, w związku z czym za przy-

¹⁴ Jako potwierdzenie swojego stanowiska, choć nie wiadomo, której z dwóch sformułowanych tu tez, Goldthorpe przytacza bogate odniesienia do literatury; por. zwłaszcza (Coleman 1986; Abbott 1992; Lindenberg i Frey 1993).

czynę uznaje się takie zjawiska, które mogą być poddane manipulacji (Goldthorpe 2001, 4).¹⁵

Z tym podejściem jest związana „mocniejsza interpretacja” przyczynowości niż w statystycznym podejściu Grangera.¹⁶ W podejściu statystycznym jedna zmienna może być „istotna przyczynowo” w stosunku do drugiej co najwyżej tymczasowo, gdyż nie można zagwarantować, że kontrolowano wszystkie istotne zmienne. W dowolnym momencie może więc pojawić się informacja, że zależność statystyczna między tymi zmiennymi nie jest odporna, a tym samym, że – przynajmniej w pewnej grupie przypadków – zależność przyczynowa między tymi zmiennymi okaże się pozorna. Problem pozorności nie dotyczy natomiast odpowiednio zaplanowanych i przeprowadzonych eksperymentów, gdyż losowe przypisanie jednostek do grupy poddanej działaniu i kontrolnej powinno wykluczyć systematyczne oddziaływanie jakiegokolwiek czynnika poza domniemaną przyczyną.

Posłużmy się przykładem Coxa i Wermuth. Niech R będzie zmienną binarną, która przyjmuje wyłącznie wartości wystąpienie-brak. W przypadku relacji monotonicznych wystąpienie R jest przyczyną wzrostu wartości K , jeśli indywiduum z cechą R ma wyższą wartość K , niż miałoby – *ceteris paribus* – w przypadku, gdyby u tego samego indywiduum wystąpił brak R . Innymi słowy, R jest przyczyną K , jeśli – przy zastosowaniu odpowiedniej kontroli – manipulacja (czy interwencja) na zmienną R daje systematyczny efekt na zmiennej K .

W modelach przyczynowych w sensie właściwym występują różne kryteria przyczynowości. Jedną grupę stanowią kryteria związane z modelami niegraficznymi (RCM oraz SEM), drugą – kryteria występujące w modelach graficznych (BN). Omówię je szczegółowo w pozostałych punktach tego rozdziału.

Obok podejścia statystycznego i przyczynowego w modelowaniu zależności przyczynowych Cox i Wermuth wyróżniają własne podejście. Aby określić, czy zachodzi zależność przyczynowa, wymagane jest określenie *procesu generującego* obserwowane zależności statystyczne między danymi zjawiskami, jak R i K . Często, jak w cytowanej wyżej krytyce podejścia statystycznego podkreślił Goldthorpe, proces generujący sformułowany będzie na innym – niższym – poziomie analizy niż stwierdzona zależność

¹⁵ Cox i Wermuth genezy współczesnych zastosowań tego podejścia w badaniach upatrują w pracach H. Simona oraz D. B. Rubina. Jest ono uogólnieniem prac J. Neymana i R. A. Fishera dotyczących planowania eksperymentów na badania nieeksperymentalne.

¹⁶ (Goldthorpe 2001, 38); por. też (Holland 1986a; 1986b; Cox 1992; Sobel 1995, 1996).

przyczynowa. Zrozumienie tego procesu traktuje się jako założenie, które jest pochodną przyjętej teorii. Jak stanie się to jasne dzięki podanej niżej szczegółowej analizie różnych odmian podejścia do modelowania przyczynowego w sensie właściwym, a także dzięki argumentacji w Zakończeniu, podanie procesu generującego nie prowadzi do sformułowania kryterium przyczynowości, które wykraczałoby poza zreferowane podejścia.

4.2 Modele Neymana-Rubina (RCM)

Zasadniczym elementem modelu RCM jest reprezentacja możliwych efektów działań jednego czynnika na daną jednostkę w porównaniu z działaniem innego czynnika na tę samą jednostkę. Ponieważ nie mogą zostać zaobserwowane obydwa, a tylko jeden, więc dla reprezentowania takich efektów wprowadza się specjalny rodzaj zmiennej, określanej za J. Neymanem zmienną potencjalną (lub kontrfaktyczną).¹⁷

Należy podkreślić odmienność tego podejścia do określenia kryterium przyczynowości zarówno w stosunku do potocznego wyobrażenia na ten temat oraz do dyskusji filozoficznych:

Ten cel, którym jest poznanie względnych skutków przyczynowych różnych manipulacji, kontrastuje z celem większości niestatystycznych dyskusji zależności przyczynowo-skutkowej, którym jest oszacowanie (lub nawet determinacja) jednej manipulacji – *jedynej* przyczyny – danego rezultatu. (Rubin 1990a, 280).

W uzasadnieniu Rubin podkreśla, że

W każdej stosunkowo skomplikowanej sytuacji w świecie rzeczywistym – czy to będzie chodziło o rośliny, zwierzęta, ludzi, samoloty czy reaktory atomowe – jestem przekonany, że na ogół nie jest możliwe przyjrzenie się zaobserwowanemu rezultatowi i realistyczne znalezienie jego *jedynej* przyczyny. (1990a, 280)

W podejściu RCM mamy jednak do czynienia z „fundamentalnym problemem wnioskowania przyczynowego” (Holland 1986b), który polega na tym, że nie jest możliwe, aby dana jednostka była i jednocześnie nie była poddana danemu działaniu. Zarejestrowane może być zajście tylko jednej z dwóch możliwości, wobec czego druga jest „kontrfaktyczna”. W modelach RCM zależność przyczynową określa się przez porównanie wartości zmien-

¹⁷ Ogólne stanowisko Neymana w kwestii modelowania zależności referuje Mortimer (1982, 178-79).

nej zmierzonej z wartością zmiennej kontrfaktycznej dla danej jednostki (wtórnie dla populacji).¹⁸

Przypomnę teraz krótko pierwotny zamysł Neymana (1923/1990). Wprowadził on pojęcie potencjalnej zmiennej, które następnie Rubin rozwinął i zastosował w badaniach zależności przyczynowych, także nieeksperymentalnych.¹⁹ Neyman opisuje eksperyment służący określeniu wydajności pola podzielonego na m poletek i obsianego v odmianami roślin. Wprowadzona przez Neymana zmienna U_{ik} reprezentuje „potencjalne zbiory”, gdyż i jest indeksem wszystkich odmian wysianych na k poletkach, a na każdym poletku aktualnie rośnie tylko jedna odmiana. Zbiór potencjalnych zmiennych U (dla wszystkich odmian w zbiorze v i wszystkich poletek w zbiorze m) jest traktowany przez Neymana jako stały, ale nieznan (Rubin 1990b, 473). Najlepszym estymatorem plonów odmiany i na danym polu jest średni potencjalny wynik dla wszystkich m poletek:

$$a_i = \sum_{k=1}^m U_{ik}/m \quad (4.1)$$

Dla określenia, jaki rodzaj roślin zostanie wysiany na danym poletku, Neyman wprowadza model urn, który – jak podkreśla Rubin – jest „stochastycznie identyczny z kompletnie zrandomizowanym eksperymentem, gdzie $n = m/v$ poletek poddano działaniu każdej z odmian” (1990b, 473). Ponieważ jednak w tym modelu urn nie można zwracać kul do urny (nie można danego pola obsiać kilkoma odmianami w tym samym sezonie), brak jest niezależności poszczególnych losowań.

Następnie Neyman wyprowadza wartości oczekiwane $E[-|U]$ i wariancji $V[-|U]$ dla powtarzanych losowań w tym eksperymencie, gdzie jedynym źródłem losowości jest losowy rozkład odmian na poletkach (gdzie U ma stałą wartość). Jeśli założyć, że x_i jest średnią wartością dla n poletek obsianych odmianą i , oraz analogicznie dla x_j , to – jak wykazuje Neyman:

$$E[x_i - x_j | U] = a_i - a_j. \quad (4.2)$$

¹⁸ Radykalnie takie podejście krytykuje A. Dawid (2000).

¹⁹ Jednym z nowszych i nagłośnionych publicznie są badania nad szczepionką przeciw wąglikowi; por. (Fellows i in. 2001; Rubin 2005).

Pokazuje więc, że standardowy estymator skutku odmiany i w porównaniu z j nie jest obciążony dla $a_i - a_j$, czyli dla wynikającej z przyjętej tu potencjalnej zmiennej oszacowania (ang. estimand) średniego skutku odmiany i w zestawieniu z j dla wszystkich poletek m .

Z kolei Neyman wykazuje, że różnica wariancji dla tych dwóch odmian jest następująca:

$$V[x_i - x_j | \mathbf{U}] = E[s_i^2/n - s_j^2/n | \mathbf{U}] - \frac{1}{m} S_{(i-j)}^2, \quad (4.3)$$

gdzie s_i^2 jest wariancją w próbce n zaobserwowanych plonów dla odmiany i , $S_{(i-j)}^2$ jest wariancją m różnic $U_{ik} - U_{jk}$.

Rubin, komentując ten wynik, podkreśla, że zwykły estymator wariancji estymacji $s_i^2/n - s_j^2/n$ jest obciążony pozytywnie dla $S_{(i-j)}^2$ różnego od 0 (dla przypadków, gdy różnica efektów poszczególnych odmian się sumuje). Ogólnie więc $S_{(i-j)}^2$ zależy od nieznanej i niezaobserwowanej korelacji między U_{ik} a U_{jk} .

Rubin w swoim komentarzu do tekstu Neymana (1923/1990) podkreśla, że najbardziej interesującym w nim elementem jest wprowadzenie potencjalnej zmiennej U_{ik} , reprezentującej plon poletka k , gdyby zostało obsiane odmianą i , której wartość jest obliczana zgodnie z opisanym przez Neymana modelem urn. Dzięki wprowadzeniu zmiennej potencjalnej można zdefiniować efekty przyczynowe w postaci $U_{ik} - U_{jk}$ oraz przyczynowe oszacowanie $a_i - a_j$ bez odwoływania się do żadnego szczegółowego modelu probabilistycznego danych (1990b, 474).

Nieadekwatność przyjętej przez Neymana reprezentacji, na jaką zwraca uwagę Rubin, dotyczy rozbieżności między zaobserwowanymi wartościami (gdzie należy uwzględnić istnienie błędów, np. pomiaru) a przyjętymi idealnymi wartościami. W późniejszych pracach (1934; 1935) Neyman próbował podać rozwiązanie tego problemu, odwołując się do hipotetycznych powtórzeń danego eksperymentu, dla których średnia „błędów technicznych” byłaby zerowa. Ze względu na to, że takie powtórzenia są czysto kontrfaktyczne, Rubin uznał propozycję Neymana za niewystarczającą i niedającą żadnych praktycznych wskazówek. Zmodyfikował więc odpowiednio przyjętą przez Neymana notację.

Dla scharakteryzowania potencjalnych plonów, uwzględniającego ewentualne wzajemne oddziaływanie na siebie poszczególnych poletek, Rubin wprowadził zbiór zmiennych $\mathbf{W} = (W_1, \dots, W_m)$, reprezentujących odmiany roślin dla m poletek. Następnie wprowadził zmienną potencjalną w postaci $Y_k(\mathbf{W})$, reprezentującą plon z poletka k , gdy wszystkie poletka obsiano zgodnie z wartościami zmiennych w \mathbf{W} . Jak podkreślił Rubin, przed faktycznym obsianiem poszczególnych poletek wartości zmiennej $Y_k(\mathbf{W})$ dla możliwych wartości \mathbf{W} są „potencjalnie obserwowalne” i żadna wartość tej zmiennej nie jest z góry przesądzona jako kontrfaktyczna. Notacja przyjęta przez Neymana jest więc szczególnym przypadkiem notacji Rubina, przy założeniu, że nie ma oddziaływania między poszczególnymi poletkami, a więc gdy plon z poletka k obsianego odmianą i jest niezależny od tego, jakie odmiany wystąpią na innych poletkach. Uogólnienie następuje przez zastąpienie w przykładzie podanym przez Neymana słów „odmiana” przez „oddziaływanie” (ang. treatment) oraz „poletko” przez „jednostka” (ang. unit) (Rubin 1990b, 475).

W podobny sposób Rubin uwzględnia w tej reprezentacji różnice między skutecznością nominalnie identycznych oddziaływań. Wprowadza kolejny zbiór zmiennych, które reprezentują poszczególne wersje oddziaływań $\mathbf{V} = (V_1, \dots, V_m)$ dla poszczególnych poletek (jednostek). Potencjalne rezultaty, uwzględniające związki między jednostkami oraz różnice w poszczególnych oddziaływaniach, reprezentuje zmienna $Y_k(\mathbf{W}, \mathbf{V})$.

Za zasadniczą zaletę tej notacji Rubin uznaje to, że „bezpośrednio odnosi się ona do zagadnień planowania [badań]” (1990b, 475).

Ponadto,

Dalszym ograniczeniem pierwotnego sformułowania Neymana jest to, że jest całkowicie związane z ocenami opartymi na randomizacji i dlatego przez pół wieku nie uważano jej za istotną dla określania skutków przyczynowych w studiach obserwacyjnych (tj. nieeksperymentalnych). (Rubin 1990b, 475)

Opracowana przez Rubina modyfikacja formalizmu Neymana stosuje się nie tylko do badań eksperymentalnych, lecz również nieeksperymentalnych. W tym celu, jak podkreśla Rubin,

wymagane jest jawne określenie modelu mechanizmu przypisania, czyli specyfikacja modelu dostarczającego prawdopodobieństwa przypisania do grupy poddanej oddziaływaniu \mathbf{W} warunkowo względem \mathbf{U} (oraz innych istotnych zmiennych), $\Pr(\mathbf{W}|\mathbf{U})$. (Rubin 1990, 476)

Taki model otrzymuje się, traktując obserwację składowych potencjalnych wyników \mathbf{U} jako proces z brakującymi danymi, gdyż w tym przypadku

$\Pr(\mathbf{W}|\mathbf{U})$ może zależeć od niezabserwowanych wartości \mathbf{U} . Eksperyment z kolei traktuje się jako proces z zaniedbywalnym mechanizmem brakujących danych, gdyż $\Pr(\mathbf{W}|\mathbf{U})$ nie zależy od \mathbf{U} . Mechanizm przypisania może więc odzwierciedlać wybieranie jednostek do próby ze skończonej populacji oraz przypisanie oddziaływania na jednostki w danym badaniu. Za zaletę tego modelu Rubin uznaje to, że

rola losowego pobierania jednostek do próbki oraz losowe przypisanie oddziaływań są jawnie reprezentowane zarówno dla zwolenników statystyki częstościowej (opartej na randomizacji oraz na modelu), jak i dla bayesianistów (łącznie ze zwolennikami bezpośredniej wiarygodności).²⁰ (Rubin 1990b, 476)

W konsekwencji zaproponowana przez Rubina – i omówiona wcześniej – miara oszacowania przyczynowego jest definiowana bez odwoływania się do określonego modelu danych $\Pr(\mathbf{U})$ czy do specyficznego mechanizmu przypisania $\Pr(\mathbf{W}|\mathbf{U})$.²¹

Zasadniczą charakterystyką takiego statystycznego podejścia do badania skutków przyczynowych jest porównanie co najmniej dwóch różnych oddziaływań na daną populację jednostek, z których każda w zasadzie mogłaby być poddana dowolnemu z tych oddziaływań.

Przejrzyste i odniesione do szerszego kontekstu, w tym dyskusji filozoficznych na temat przyczynowości, omówienie modelu RCM podaje (Holland 1986b). Model RCM jest określony dla danej populacji, czyli pewnego uniwersum jednostek U . Jednostki u w U (np. ludzie, gospodarstwa domowe, sprzęt laboratoryjny, pole) stanowią podstawowy przedmiot badania. Z kolei dla jednostek określa się zmienną jako funkcję rzeczywistą zdefiniowaną dla wszystkich u w U , a której wartość dla danej jednostki u jest liczbą „przypisaną przez pewien proces pomiaru” tej jednostce. Szczególnego rodzaju zmienną jest tzw. zmienna odpowiedzi, czyli zmienna wyjaś-

²⁰ W tym przypadku chodzi o nieliczną grupę statystyków, którzy nie są ani zwolennikami klasycznej statystyki częstościowej ani bayesianistami, a za jedyną dopuszczalną miarę informacji, jakich dostarcza nam doświadczenie, uznają Fishera iloraz wiarygodności, odrzucając częstościowy model danych $\Pr(\mathbf{U})$ oraz bayesowskie prawdopodobieństwo pierwotne $\Pr(\mathbf{W})$, jako arbitralne; por. (Berger 1998).

²¹ Rubin podkreśla także, że opracowanie przez niego w końcu lat 1970. modyfikacji notacji Neymana, którą można zastosować także do badań nieeksperymentalnych, zaowocowało licznymi aplikacjami, których przykładem są (Robins 1987; Rosenbaum 1987; Greenland i Poole 1988; Holland 1988; Pratt i Schlaifer 1988; Smith i Sudgen 1988; Heckman 1989; Holland 1989; Sudgen 1989; Rubin 1990a; Sobel 1990; Dempster 1991; Efron i Feldman 1991; Gelman i King 1991; Rubin 1991).

niana, $Y(u)$, która posiada także wartość dla każdej jednostki w U oraz jest przedmiotem zainteresowania badacza, zmierzającego do wyjaśnienia różnic między wartościami $Y(u)$ dla poszczególnych jednostek w u . Takie wyjaśnienie odwołuje się do innej zmiennej, jaką jest atrybut jednostek w U , oznaczany jako $A(u)$.

Holland proponuje, by przez prawdopodobieństwo rozumieć proporcję jednostek w U , a przez wartość oczekiwaną danej zmiennej, jej wartość średnią dla wszystkich jednostek w U . Z kolei warunkowa wartość oczekiwana jest także wartością średnią dla podzbioru jednostek wyznaczonego przez odpowiednie wartości zmiennych, na których się warunkuje (Holland 1986b, 945). Rozkład łączny Y oraz A w U jest określony przez $\Pr(Y = y, A = a)$ i odczytywany jako stosunek tych u w U do wszystkich jednostek w U , dla których są spełnione warunki $Y(u) = y$ oraz $A(u) = a$. Rozkład łączny pozwala na określenie wartości parametrów takich, jak np. rozkład warunkowy $\Pr(Y = y | A = a) = \Pr(Y = y, A = a) / \Pr(A = a)$, opisujący, jak zmienia się rozkład wartości Y wraz ze zmianą A , oraz regresja Y na A rozumiana tu jako warunkowa wartość oczekiwana $E(Y | A = a)$.

Wnioskowanie asocjacyjne, które Holland przeciwstawia wnioskowaniu przyczynowemu, polega w tym przypadku na statystycznym wnioskowaniu (estymacja, testy, rozkład wtórny) parametrów wiążących Y oraz A na podstawie zebranych danych dotyczących jednostek w U .

Model RCM, za Neymanem, jako najprostszą sytuację pozwalającą określić zależności przyczynowe, przyjmuje eksperyment, jak krótko wyraża to Holland:

nie sędzę, że eksperyment jest *jedynym* kontekstem odpowiednim do dyskusowania przyczynowości, ale jestem przekonany, że jest on *najprostszym* takim kontekstem. (1986b, 946)

Podobnie więc jak robił to Neyman, lecz upraszczając w sposób zaproponowany przez Rubina, wprowadza się w modelu RCM reprezentację oddziaływania, które Holland traktuje jako zamiennie z przyczyną:

Należy podkreślić, że używając zamiennie terminów przyczyna i oddziaływanie, nie zamierzam ograniczać dyskusji wyłącznie do aktywności w ramach kontrolowanego badania z randomizacją, lecz by podkreślić ideę, która nie jest dostrzegana w ogólnych dyskusjach o przyczynowości. Chodzi o to, że skutek danej przyczyny jest *zawsze* względny w stosunku do innej przyczyny. (Holland 1986b, 946)

Względność polega na tym, że aby wskazać taki skutek, należy porównać dwie grupy: jedną poddaną oddziaływaniu, a drugą (kontrolną) nie. Dla każ-

dej z tych grup następnie, jak twierdzi Holland, należy przyjąć, że działa inna przyczyna. Dlatego też kluczowego znaczenia nabiera *potencjalność* poddania danej jednostki oddziaływaniu jednej lub drugiej przyczyny (ang. potential for exposing or not exposing). Upraszczając, można założyć, że są tylko dwie przyczyny, reprezentowane jako t oddziaływanie oraz c kontrola. Rodzaj oddziaływania, jakiemu poddano daną jednostkę, reprezentuje zmienna S , gdzie $S = t$ w przypadku oddziaływania na jednostkę oraz $S = c$, gdy jednostka jest w grupie kontrolnej. W przypadku badań eksperymentalnych (kontrolowanych) S jest skonstruowane przez badacza, natomiast w przypadku badań nieeksperymentalnych S jest do pewnego stopnia zdeterminowane przez czynniki pozostające poza kontrolą badacza (Holland 1986b, 946). Zasadnicze dla modelu RCM jest to, że wartość $S(u)$ mogłaby być inna dla każdej jednostki w U .

Mimo więc podobieństwa do zmiennej $A(u)$ reprezentującej atrybut u , $S(u)$, reprezentująca poddanie oddziaływaniu określonej przyczyny, różni się przede wszystkim przez to, że może potencjalnie mieć inną wartość, podczas gdy to nie jest dopuszczalne w przypadku atrybutu danej jednostki. Ponadto, wartość zmiennej $S(u)$ określa, która z możliwych wartości: Y_t czy Y_c , została zaobserwowana dla określonej jednostki. W ogólnej postaci zaobserwowana wartość odpowiedzi u na oddziaływanie, jakiemu została poddana, to $Y_{S(u)}(u)$, a zmienna reprezentująca zaobserwowaną odpowiedź to Y_S . Mimo więc, że model zawiera trzy zmienne: S , Y_t oraz Y_c , to obserwacja dotyczy tylko dwóch: S oraz Y_S .

Odpowiednio też wymagana jest modyfikacja zmiennej Y , która ma reprezentować zmianę, jaka dokonuje się w danej jednostce pod wpływem określonego oddziaływania. Ponieważ może być ona poddana różnym oddziaływaniom, więc zmienna Y musi to reprezentować. W standardowej wersji modelu RCM tę reprezentację wprowadza się poprzez odpowiedni indeks tej zmiennej, korespondujący z rodzajem oddziaływania, jakie jest rozważane w danym badaniu, np. Y_t oraz Y_c . Dla danej jednostki u interpretacja wartości zmiennej uwzględniającej oddziaływanie, jakiemu może być poddana ta jednostka, np. $Y_t(u)$, jest następująca: jest to wartość odpowiedzi, którą zaobserwowano by, gdyby jednostkę u poddano oddziaływaniu t . Podobnie dla $Y_c(u)$ – to jest wartość odpowiedzi, którą zaobserwowano by, gdyby tę samą jednostkę u poddano oddziaływaniu c .

Zgodnie z przywołanym wcześniej komentarzem Rubina oraz Hollanda, za miarę efektu przyczynowego uznaje się więc różnicę między wartością odpowiedzi danej jednostki na jedno oraz drugie oddziaływanie:

$$Y_t(u) - Y_c(u), \quad (4.4)$$

czyli – jak proponuje interpretować tę różnicę Holland – efekt przyczynowy oddziaływania t na daną jednostkę u (względem oddziaływania c). Zgodnie więc z podejściem RCM do modelowania zależności przyczynowych „Wnioskowanie przyczynowe dotyczy ostatecznie efektów przyczyn dla poszczególnych jednostek” (Holland 1986b, 947).

Pomijając warstwę interpretacyjną, można podać czysto matematyczną postać modelu RCM następująco. Model RCM = (U, K, Y, S) , gdzie U oraz K są zbiorami, Y jest funkcją rzeczywistą zdefiniowaną na $U \times K$, a S jest odwzorowaniem zbioru U w zbiór K . U jest populacją jednostek, K jest zbiorem opisów różnych przyczyn czy oddziaływań uwzględnianych w danym badaniu. Dla dowolnego $u \in U$ oraz $k \in K$ $Y(u, k)$ jest wartością odpowiedzi, którą zmierzono by na u , gdyby u była poddana oddziaływaniu k . Wartością $S(u)$ jest oddziaływanie, jakiemu faktycznie poddano u przed zmierzeniem odpowiedzi u . W omawianym wcześniej przypadku wykorzystywano fakt, że $Y_k(u) = Y(u, k)$ oraz $K = \{t, c\}$. Z kolei atrybutem jest taka zmienna, której wartość nie zmienia się pod wpływem poddania u danemu oddziaływaniu: $X(u, k) = X(u)$.

Efekt oddziaływania t na u względem oddziaływania c na tę samą jednostkę wyraża podana wcześniej różnica wartości zmiennej Y dla tych oddziaływań na u . Natomiast zaobserwowaną odpowiedzią u na dane oddziaływanie jest $Y_S(u) = Y(u, S(u))$.

Hipotezę o braku zależności przyczynowej (hipoteza zerowa Fishera) można – za Rubinem – w tym formalizmie wyrazić jako $Y(u, k) = Y(u)$. Holland proponuje, by analogicznie wyrazić tezę o homogeniczności jednostek: $Y(u, k) = Y(k)$.

Przy takim podejściu do modelowania przyczynowości, jak zauważa Holland, zasadniczy problem dotyczy tego, że pewne z wartości zmiennej Y dla danej jednostki nie są obserwowalne. Ten właśnie problem Holland określa jako fundamentalny problem wnioskowań przyczynowych:

Niemożliwe jest zaobserwowanie wartości $Y_t(u)$ oraz $Y_c(u)$ dla tej samej jednostki, a zatem, niemożliwe jest zaobserwowanie efektu t na u . (1986b, 947).

Holland podkreśla jednak, że z tego problemu nie wynika bezpośrednio niemożliwość wyciągnięcia jakichkolwiek wniosków przyczynowych (globalny sceptycyzm przyczynowy). Jak podkreśla bowiem, z tego, że jest niemożliwe zaobserwowanie obu tych wartości (odpowiedzi danej jednostki na dwa różne oddziaływania) nie wynika, że nie ma jakichkolwiek istotnych

informacji o niezaobserwowanej wartości. Możliwe rozwiązania tego problemu Holland grupuje w dwie kategorie: rozwiązania naukowe oraz rozwiązania statystyczne.

Rozwiązanie naukowe odwołuje się do różnych (nietestowalnych) założeń o homogeniczności lub niezmienniczości. Ilustracją takiego postępowania może być badanie aparatury laboratoryjnej. Zakładając, że zmierzona przed danym eksperymentem wartość $Y_c(u)$ nie uległa zmianie w czasie eksperymentu, i poddając badaną aparaturę oddziaływaniu t oraz mierząc $Y_t(u)$, naukowiec jest w stanie określić (względny) skutek przyczynowy oddziaływania t . Jednak takie rozwiązanie fundamentalnego problemu wnioskowania przyczynowego oparte jest na nietestowalnym założeniu o jednorodności badanej aparatury laboratoryjnej. Jednak

Przez uważną pracę naukowiec może przekonać siebie i innych, że to założenie jest słuszne, chociaż nie może mieć absolutnej pewności. Nauka posunęła się bardzo daleko, wykorzystując takie podejście. To naukowe rozwiązanie jest także powszedniością naszego życia codziennego. (Holland 1986b, 947)

Inne jest rozwiązanie statystyczne. Wprowadza się pojęcie *średniego efektu przyczynowego* T danego oddziaływania t (względem c) w populacji U jako wartości oczekiwanej różnicy $Y_t(u) - Y_c(u)$ dla jednostek u w U :

$$E(Y_t - Y_c) = T. \quad (4.5)$$

Korzystając z teorii prawdopodobieństwa, można dalej pokazać, że

$$T = E(Y_t) - E(Y_c). \quad (4.6)$$

Takie z kolei wyrażenie T pozwala na wyprowadzenie wartości średniego efektu przyczynowego z odwołaniem się do pomiarów dokonywanych na różnych jednostkach, np. te poddane oddziaływaniu t można wykorzystać do określenia średniej wartości Y_t w U , czyli $E(Y_t)$. Podobnie dla oddziaływania c w tej samej populacji. Jak podkreśla Holland,

Dokładny sposób dobierania jednostek poddanych oddziaływaniu t lub c jest bardzo ważny i prowadzi do zagadnień zwykle rozważanych przy dobrym statystycznym planowaniu eksperymentów. (1986b, 947)

Należy podkreślić, że rozwiązanie statystyczne zastępuje niemożliwy do *zaobserwowania* efekt (względny) oddziaływania t na daną jednostkę, przez możliwy do *estymacji* średni efekt (względny) oddziaływania t w danej populacji jednostek.

O tej populacji w rozwiązaniu statystycznym zakłada się, że jest „duża”, a zaobserwowane dane dla każdej jednostki są wartościami pary zmiennych (Y_S, S) . Z tych danych można uzyskać jedynie informacje o warunkowych wartościach oczekiwanych poszczególnych oddziaływań:

$$E(Y_S | S = t) = E(Y_t | S = t) \quad (4.7)$$

$$E(Y_S | S = c) = E(Y_c | S = c). \quad (4.8)$$

Te warunkowe wartości oczekiwane nie są dokładnie tym samym, co bezwarunkowe $E(Y_t)$ oraz $E(Y_c)$. I tak, $E(Y_t)$ jest średnią wartością $Y_t(u)$ dla wszystkich u w populacji U , natomiast $E(Y_t | S = t)$ jest średnią wartością $Y_t(u)$ dla tych u , które były poddane oddziaływaniu t . W ogólnym przypadku te wartości mogą być rozbieżne, np. gdy $S(u) = t$ dla wszystkich jednostek, dla których wartość $Y_t(u)$ jest mała, to $E(Y_t | S = t)$ będzie mniejsze od $E(Y_t)$.

Innym problemem rozwiązania statystycznego, na jaki zwraca uwagę Holland, jest to, że średni efekt przyczynowy może nie reprezentować wartości, która przybliżałaby wartości efektu przyczynowego dla określonej jednostki. Taka sytuacja wystąpi szczególnie wyraziście w przypadku, gdy jest duża wariancja efektów przyczynowych $Y_t(u) - Y_c(u)$ w danej populacji. Wówczas nawet przy bardzo uważnej i dokładnej estymacji T będzie całkowicie nieistotne przy określeniu wartości efektu przyczynowego dla danej jednostki w tej populacji.

Rozwiązanie naukowe można, jak pokazuje Holland (1986b, 948-49), zastosować, dodając do modelu RCM dodatkowe założenia. Takimi dwoma założeniami stosowanymi często w odniesieniu do urządzeń fizycznych, zarówno w badaniach naukowych, jak i życiu codziennym, są założenie stabilności w czasie oraz założenie krótkotrwałości efektu. Pierwsze określa, że wartość $Y_c(u)$ nie zależy od tego, kiedy dokonano się oddziaływanie c na u oraz zmierzenie Y_c . Drugie określa, że wartość $Y_t(u)$ nie zależy od tego, czy wcześniej zastosowano oddziaływanie c do u – oddziaływanie c oraz jego pomiar są na tyle krótkotrwałe, że nie wpływają na późniejsze oddziaływanie t na u i jego pomiar.

Inny sposób wprowadzenia rozwiązania naukowego w modelu RCM to dodanie założenia o homogeniczności (jednorodności) jednostek. Uzyskuje się je, mierząc (względny) efekt przyczynowy t poprzez różnicę tego oddziaływania na określonej jednostkę u_1 w stosunku do oddziaływania c na inną jednostkę u_2 : $Y_t(u_1) - Y_c(u_2)$. O tych jednostkach zakłada się ponadto, że

są jednorodne: $Y_t(u_1) = Y_t(u_2)$ oraz $Y_c(u_1) = Y_c(u_2)$. Jak podkreśla Holland, nie można uzyskać całkowitej pewności co do słuszności tego założenia w konkretnym przypadku, jednak można nadać mu miano wiarygodności, badając uważnie, czy poszczególne jednostki „wyglądają” tak samo pod istotnymi względami (1986b, 948).

Często stosowanym sposobem wykorzystania rozwiązania statystycznego jest zastosowanie randomizacji przy wyborze jednostek poddanych obserwacji w danym badaniu (założenie o niezależności). Pozwała ono na wykazanie, że różne w ogólnym przypadku wartości $E(Y_t | S = t)$ oraz $E(Y_t)$ są równe. Zgodnie z założeniem o randomizacji, jednostki są poddane oddziaływaniu t lub c zgodnie z pewnym losowym procesem, który decyduje o wyborze oddziaływania dla każdej z jednostek. A zatem, S , czyli to, któremu z oddziaływań poddane jest u , jest statystycznie *niezależne* od wszystkich innych zmiennych, w tym Y_t oraz Y_c . Przy założeniu randomizacji średni efekt przyczynowy spełnia więc równanie:

$$T = E(Y_S | S = t) - E(Y_S | S = c). \quad (4.9)$$

Można więc wykorzystać zaobserwowane dane (S, Y_S) do estymacji T , odejmując różnicę wartości Y_S dla jednostek, u których zaobserwowano $S = t$, w stosunku do wartości Y_S dla jednostek, u których zaobserwowano $S = c$.²²

A zatem, jeśli możliwa jest randomizacja, wówczas zawsze możliwe jest estymowanie średniego efektu przyczynowego danego oddziaływania w stosunku do innego. Dla dużych U taka estymacja może być przeprowadzona z dużą dokładnością.

Inne założenie umożliwiające rozwiązanie statystyczne to założenie stałego efektu.²³ Przyjmuje się, że dane oddziaływanie ma ten sam efekt dla wszystkich jednostek danej populacji:

$$T = Y_t(u) - Y_c(u) \text{ dla każdego } u \text{ w } U. \quad (4.10)$$

²² Wyrażoną w powyższym równaniu różnicę Holland określa jako średni efekt przyczynowy *prima facie* (powierzchnowy) i podkreśla, że bez założenia o niezależności nie jest on równy prawdziwemu efektowi przyczynowemu wyrażonemu przez $E(Y_t - Y_c) = T$ (1986b, 949).

²³ Innym określeniem tego założenia, jakie można spotkać w literaturze statystycznej, jest założenie addytywności, gdyż oddziaływanie t dodaje stałą wartość T do odpowiedzi w grupie kontrolnej dla każdej jednostki.

Przy tym założeniu T jest efektem przyczynowym dla każdej jednostki i unika się zarysowanego wcześniej problemu rozbieżności między efektem przyczynowym oddziaływania t na daną jednostkę a estymowanym średnim efektem przyczynowym T w danej populacji. T – przy założeniu stałego efektu – pozwala na dokonywanie wnioskowań przyczynowych na poziomie poszczególnych jednostek.

Można także założenie stałego efektu poddać sprawdzeniu. Populację U dzieli się na podpopulacje U_1, \dots, U_n i dla każdej z nich osobno estymuje się średni efekt przyczynowy T_1, \dots, T_n . Jeśli wartości poszczególnych średnich efektów przyczynowych są różne, wówczas nie można przyjąć założenia o stałym efekcie przyczynowym. Jeśli natomiast nie są różne, wówczas założenie stałego efektu „może być wiarygodne” (Holland 1986b, 949).

Założenie stałego efektu wynika z założenia homogeniczności. Jeśli bowiem $Y_t(u_1) = Y_t(u_2)$ oraz $Y_c(u_1) = Y_c(u_2)$, to $Y_t(u_1) - Y_c(u_1) = Y_t(u_2) - Y_c(u_2)$. Założenie stałego efektu jest więc osłabieniem założenia homogeniczności jednostek danej populacji, gdyż nie implikuje równości między średnim efektem przyczynowym a *prima facie* efektem przyczynowym (Holland 1986b, 949).

W odniesieniu do badań nieeksperymentalnych Holland podkreśla, że wykorzystuje się tam zmienne przed oddziaływaniem (ang. pre-exposure), aby w miejsce założenia o niezależności wprowadzić mniej restryktywne założenie o *niezależności warunkowej*. Takie właśnie założenie – zwane tu *mocną zaniedbywalnością* ze względu na współzmiennne Z – jest stosowane w aplikacjach modelu RCM do badań nieeksperymentalnych (Rosenbaum i Rubin 1983, 42):

$$(Y_t, Y_c) \perp\!\!\!\perp S \mid Z \text{ oraz } 0 < \Pr(S = t \mid Z) < 1, \quad (4.11)$$

gdzie Z to zbiór zmiennych sprzed poddania oddziaływaniu, a symbol $\perp\!\!\!\perp$ reprezentuje niezależność probabilistyczną (Dawid 1979). Jeśli ten warunek jest spełniony dla takich Z , które są zaobserwowane, wówczas mówi się o *mocnej zaniedbywalności* (bez dodatkowej kwalifikacji).

Wprowadźmy za Rosenbaumem i Rubinem pojęcie *wyniku wyważonego* $b(Z)$ jako takiej funkcji zaobserwowanych współzmiennnych Z , że warunkowy rozkład Z (względem $b(Z)$) jest taki sam dla jednostek poddanych oddziaływaniu i poddanych kontroli:

$$S \perp\!\!\!\perp Z \mid b(Z).^{24} \quad (4.12)$$

Kolejnym ważnym pojęciem jest prawdopodobieństwo warunkowe przypisania do danego rodzaju oddziaływania, określane przez Rosenbauma i Rubina jako *wynik skłonności*:²⁵

$$e(Z) = \Pr(S = t \mid Z). \quad (4.13)$$

O tym prawdopodobieństwie zakłada się, że ma rozkład wielomianowy dla poszczególnych wartości S oraz Z (czyli, że jest niezależne dla poszczególnych rodzajów oddziaływań).

Wymogiem w stosunku do wyniku skłonności jest, aby był bardziej zgrubny niż wynik wyważony w podanym sensie:

$$e(Z) = f(b(Z)) \text{ dla pewnej funkcji } f. \quad (4.14)$$

Z tego wymogu wynika niezależność warunkowa zaobserwowanych współzmiennych Z oraz przypisanie do danego oddziaływania S :

$$S \perp\!\!\!\perp Z \mid e(Z). \quad (4.15)$$

To twierdzenie oznacza, że w pewnej podgrupie jednostki (lub dopasowane pary typu oddziaływanie-kontrola), które są jednorodne pod względem przypisania oddziaływania, mają taki sam rozkład zaobserwowanych współzmiennych Z (Rosenbaum, Rubin 1983, 44). W odniesieniu do badań nieeksperymentalnych jednym z najważniejszych rezultatów cytowanej pracy jest dowód twierdzenia (Rosenbaum Rubin 1983, 45), że gdy przypisanie do danego oddziaływania jest mocno zaniedbywalne, to uwzględnienie wyniku wyważonego $b(Z)$ jest wystarczające do otrzymania nietendycyjnego estymatorów średniego efektu przyczynowego:

$$(Y_t, Y_c) \perp\!\!\!\perp S \mid Z \text{ oraz } 0 < \Pr(S = t \mid Z) < 1 \quad (4.16)$$

które dla wszystkich wartości Z implikują, że

²⁴ Najprostszy wynik wyważony jest następujący $b(Z) = Z$.

²⁵ Por. też (D'Agostino i Rubin 2000; Imai i van Dyk 2004; Lunceford i Davidian 2004).

$$(Y_t, Y_c) \perp\!\!\!\perp S \mid b(Z) \text{ oraz } 0 < \Pr(S = t \mid b(Z)) < 1, \quad (4.17)$$

dla wszystkich $b(Z)$.

Konsekwencją tego rezultatu jest to, że jeśli wynik oddziaływania się nie zmieni dla dowolnej z jednostek mających zaobserwowane przed oddziaływaniem własności Z , to nie zmieni się on również, gdy uwzględni się taką funkcję Z , która zachowuje niezależność przypisanego jednostce rodzaju oddziaływania od jej własności.

Na podstawie tego rezultatu Rosenbaum i Rubin (1983, 45-46) wykazują, że nieobciążonym estymatorem średniego efektu przyczynowego t (względem c) dla danej wartości wyniku wyważonego jest różnica między zaobserwowaną średnią dla tejże wartości wyniku wyrażonego, o ile przypisanie do oddziaływania jest mocno zaniechane:

$$E(Y_t \mid b(Z), S = t) - E(Y_c \mid b(Z), S = c) = E(Y_t - Y_c \mid b(Z)), \quad (4.18)$$

gdzie $b(x)$ jest wynikiem wyważonym, a przypisanie oddziaływania jest mocno zaniechane.

Z tego twierdzenia wyprowadzają następnie wnioski, że takiego nieobciążonego estymatora średniego efektu przyczynowego dostarczają – przy założeniu mocno zaniechanego przypisania oddziaływania – dopasowanie par, podklasyfikacja oraz uwzględnienie kowariancji.

Ogólnym wnioskiem z omówionych rezultatów jest, że w badaniach obserwacyjnych uwzględnienie wyniku skłonności powinno być – jak podkreślają Rosenbaum i Rubin – istotnym składnikiem analizy. Wystąpienie bowiem resztowego obciążenia w wyniku skłonności świadczy o potencjalnym obciążeniu w estymowanym efekcie oddziaływania (Rosenbaum i Rubin 1983, 48).

Sposób wykorzystania wyniku skłonności do skorygowania zmiennych konfundujących w badaniu nieeksperymentalnym Rosenbaum i Rubin ilustrują przykładami, odwołującymi się do standardowych technik korygowania konfunderów w badaniach nieeksperymentalnych, jakimi są: dopasowanie par, podklasyfikacja oraz skorygowanie kowariancji (Rosenbaum i Rubin 1983, 48-55).

Dopasowanie to sposób pobierania próbki z licznego zasobu jednostek, które mogą znaleźć się w grupie kontrolnej, i stworzenia takiej grupy kontrolnej średnich rozmiarów, w której rozkład współzmiennych jest podobny do rozkładu w grupie oddziaływania. Przewagą dopasowania w

stosunku do innych sposobów pobierania próbek, np. skorygowanie kowariancji w próbkach, jest – zdaniem Rosenbauma i Rubina – to, że:

1) łatwiej badaczowi jest dostrzec równoważność grupy kontrolnej i grupy oddziaływania oraz przeprowadzić analizę dla dopasowanych par, która skoryguje zmienne konfundujące;

2) jest mniejsza wariancja estymatorów średniego efektu oddziaływania dla dopasowanych niż losowych próbek, gdyż rozkłady Z w grupie oddziaływania i kontrolnej są bardziej zbliżone w próbkach dopasowanych niż w losowych;

3) oparte na modelu skorygowanie dla próbek dopasowanych jest zwykle bardziej odporne na odchylenia od postaci przyjętej dla danego modelu niż dla próbek losowych z uwagi na mniejsze wykorzystywanie ekstrapolacji modelu;

4) w badaniach z dużym zasobem do pobierania jednostek do grupy kontrolnej oraz wieloma zmiennymi konfundującymi, te ostatnie można kontrolować przez wielozmiennowe dopasowanie, podczas gdy ostatecznie poddane badaniu próbki o małych rozmiarach nie pozwalają kontrolować wszystkich zmiennych za pomocą metod opartych na modelu.

W idealnym przypadku obie grupy (oddziaływania i kontrolna) powinny być dopasowane dla wszystkich współzmiennych w Z tak, aby rozkłady dla tych zmiennych w próbkach pobieranych z obu grup były identyczne. Z własności wyniku wyważonego wynika, że wystarczy dopasowanie dla dowolnej wartości takiego wyniku, aby otrzymać żądane rozkłady dla obu prób. Ponadto, z twierdzenia, że nieobciążonym estymatorem średniego efektu przyczynowego t (względem c) dla danej wartości wyniku wyważonego jest różnica między zaobserwowaną średnią dla tejże wartości wyniku wyważonego, o ile przypisanie oddziaływania jest mocno zaniedbywalne, wynika też (Rosenbaum i Rubin 1983, 46), że do takiego estymatora prowadzi dokładne dopasowanie dla danego wyniku wyważonego. Ponieważ w badaniach nie jest możliwe otrzymanie idealnego dopasowania, więc Rosenbaum i Rubin stosują dopasowanie przybliżone. Jako jego kryterium wykorzystują równą procentową redukcję obciążenia: dana metoda w równym procencie redukuje obciążenie, jeśli obciążenie dla każdej zmiennej w Z jest zredukowane o taki sam procent (Rosenbaum i Rubin 1983, 49). Metody dopasowania niespełniające tego warunku mają własność zwiększania obciążenia dla pewnej liniowej funkcji Z . Przy spełnieniu dodatkowych założeń (dopasowanie dokonuje się przed pomiarem odpowiedzi na oddziaływanie, zależność między odpowiedzią na dane oddziaływanie a współzmiennymi Z jest

w przybliżeniu liniowa) metody w równym procencie redukujące obciążenie prowadzą do różnic w średniej między poszczególnymi próbkami, która jest mniejsza niż dla próbek losowych. Rosenbaum i Rubin wykazują następnie, że przy osłabionych założeniach co do rozkładu Z dopasowanie dla wyniku skłonności danej populacji jest metodą w równym procencie redukującą obciążenie (1983, 49-51).

Kolejną ważną metodą skorygowania konfunderów jest podklasyfikacja. Polega ona na podziale grup oddziaływania i kontrolnej zgodnie ze współzmiennymi w Z . Zgodnie z dowiedzionym przez Rosenbauma i Rubina twierdzeniem (1983, 46), oczekiwana różnica w średnim oddziaływaniu dla jednostek wybranych do próby za pomocą $b(x)$, które jest stałe dla wszystkich jednostek, i przy mocno zaniechwalnym przypisaniu oddziaływania, jest równa średniemu efektowi przyczynowemu. Ponadto, średnia ważona tych różnic (czyli różnica skorygowana bezpośrednio) jest nieobciążona w stosunku do efektu oddziaływania, gdy wagi są równe stosunkowi populacji dla $b(x)$.

Jako metoda wielozmiennowego korygowania, podklasyfikacja ma tę zaletę, że wiąże się z bezpośrednim porównaniem grup jednostek dla każdej podklasy i dzięki temu może być czytelna dla odbiorcy niezaznajomionego ze statystyką. Zasadniczy jednak problem podklasyfikacji, na który zwrócili uwagę W. Cochran i S. Chambers (1965) polega na tym, że liczba podklas gwałtownie rośnie dla zwiększonej liczby konfunderów. I tak, jeśli każdy z p konfunderów miałby tylko dwie wartości, to wśród 2^p podklas wiele nie zawierałoby jednostek poddanych jednemu bądź drugiemu oddziaływaniu. „Podklasyfikacja oparta na wyniku skłonności jest naturalnym sposobem uniknięcia tej trudności”, jak podkreślają Rosenbaum i Rubin (1983, 52).

Szczegółową ilustracją tego podejścia są badania nieeksperymentalne efektów różnych sposobów leczenia choroby wieńcowej serca. Oddziaływania w tym przypadku to operacja wszczepienia by-passów ($S = t$) oraz leczenie farmakologiczne ($S = c$). Współzmiennymi w Z są pomiary kliniczne, hemodynamiczne oraz demograficzne dokonane dla każdego z pacjentów przed przypisaniem ich do jednego z oddziaływań.

Wynik skłonności estymowano za pomocą modelu logit dla S warunkowo względem Z . Wykorzystując fakt, że podklasyfikacja z pięcioma podklasami usuwa co najmniej 90% obciążenia dla wielu rozkładów ciągłych (Cochran 1968), skonstruowano pięć podklas o równej liczebności 303 osób dla kwantyli rozkładu wyniku skłonności w próbie. Mediana dla średniego

efektu przyczynowego bez podklasyfikacji wyniosła 10,9, a z podklasyfikacją 0,2 (Rosenbaum, Rubin 1983, 52).

Trzecim standardowym sposobem korygowania, do którego odnoszą się omawiani autorzy jest korygowanie współzmiennych. Estymator punktowy efektu oddziaływania otrzymany w analizie korygowania współzmiennych jest, jak podkreślają, równy estymatorowi otrzymanemu przy korygowaniu jednozmiennowej kowariancji dla próbkowego liniowego wyróżnika opartego na Z (Rosenbaum, Rubin 1983, 53-54). Warunkiem jednak usunięcia obciążenia jest, by liniowo zależny od Z wyróżnik był mocno skorelowany z wynikiem skłonności, która nie musi być funkcją liniową Z .

Krótkie zestawienie podejścia RCM z innymi poprzedzonym zwartym zestawieniem kryteriów przyczynowości, jakie się w nim proponuje. Przyjmę, jak w poprzednim punkcie, że R jest domniemaną przyczyną K . Najprostsze z kryteriów RCM odwołuje się do takiej relacji między dwoma wartościami K (przy R występującym – K_w oraz nieobecny – K_b), że w pewnej populacji badanych jednostek średnia wartość K_w przewyższa średnią wartość K_b . W takim przypadku mamy do czynienia z pojęciem średniego efektu, który jest empirycznie testowalny.

Mocniejsze kryterium polega na żądaniu, by różnica wartości K_w i K_b występowała w przypadku każdej jednostki w danej populacji. Natomiast przy jeszcze mocniejszym kryterium ta różnica jest taka sama dla wszystkich indywidualów: $K_w - K_b = \Delta$.

Te dwa ostatnie wymogi, które byłyby najbliższe deklaracjom Rubina, jakie zacytowałem na początku tego punktu, nie są jednak bezpośrednio testowalne, co podnoszone jest jako zarzut przeciw podejściu RCM. Cox i Wermuth sugerują możliwość – zachodzącą przynajmniej w niektórych przypadkach – pośredniej testowalności, która polegałaby na podzieleniu indywidualów na grupy, np. na podstawie jednej ze współzmiennych B , i zaobserwowaniu, czy dla każdego z badanych indywidualów różnica między dwoma wartościami K ma ten sam znak. Należy przy tym uwzględnić fakt, że R może mieć bardzo zróżnicowany wpływ na poszczególne indywiduala (2001, 68). Do tego problemu powrócę w Punkcie 5.1.

Można zasygnalizować także pewną teoretyczną trudność podejścia RCM, którą podziеляją i inne podejścia do modelowania przyczynowego w sensie właściwym. Mianowicie zmienne „czysto wyjaśniające”²⁶ mogą reprezentować *atrybuty wsobne* jednostek (zwane niekiedy zmiennymi

²⁶ W terminologii ekonometrycznej są to zmienne czysto egzogeniczne.

strukturalnymi), które stanowią część ich istotnej charakterystyki. Tego typu zmienne w podejściu RCM nie powinny być uważane za zmienne wchodzące w związki przyczynowe, gdyż nie można o nich zakładać – nawet czysto hipotetycznie – że mogłyby przyjąć inną wartość od zaobserwowanej bez wpływu na tożsamość jednostki. Jako przykład Cox i Wermuth (2001, 68) podają płeć, traktowaną w większości badań jako atrybut wsobny. W takim przypadku trudno sformułować spójną interpretację określonej na gruncie RCM wartości, jaką K kontrfaktycznie przyjęłoby, gdyby przyjąć, że np. określona kobieta zostałaaby – *ceteris paribus* – mężczyzną.

Na koniec rozważań poświęconych podejściu RCM warto zwrócić uwagę na jego relacje do innych podejść w modelowaniu zależności przyczynowych. Skupię się tu przede wszystkim na podejściu statystycznym w ujęciu Grangera oraz podejściu SEM.

W rekonstrukcji podejścia Grangera na gruncie RCM w zasadzie zbędne są zmienne potencjalne, co można potraktować jako potwierdzenie statystycznego, a nie przyczynowego (w sensie właściwym) charakteru tego podejścia. Taka rekonstrukcja może być przeprowadzona następująco. Niech zmienne X , Y i Z oznaczają zmienne zdefiniowane dla danej populacji. Zmienne X i Y są warunkowo niezależne, gdy spełniony jest warunek:

$$\Pr(Y=y | X=x, Z=z) = \Pr(Y=y | Z=z). \quad (4.19)$$

Niezależność warunkowa jest mocniejszą wersją stwierdzenia Grangera, że wartości zmiennej X nie mogą służyć do przewidywania wartości Y warunkowo względem korygowania wartości Z .

W podejściu Grangera odnoszącym się do szeregów czasowych wartość zmiennej Y jest określona w pewnym momencie czasu s , natomiast wartości X i Y są określone w s lub innym momencie czasowym r wcześniejszym od s . Brak zależności przyczynowej w sensie Grangera między X a Y można zdefiniować jako niezależność warunkową tych zmiennych warunkowo względem Z . Zależność przyczynowa w sensie Grangera zachodzi więc między tymi zmiennymi, jeśli różne wartości X prowadzą do przewidywania różnych predyktywnych rozkładów Y , uwzględniając zarówno informacje zawarte w X , jak i w Z , a więc gdy X pozwala przewidzieć wartości Y , nawet po uwzględnieniu Z .

Na gruncie modelu RCM zmiennej X odpowiada zmienna S , zmiennej Y – zmienna Y_S , a Z to zbiór zmiennych sprzed poddania oddziaływaniu. W tym przypadku warunkowa niezależność przedstawia się następująco:

$$\Pr(Y_S = y \mid S = t, Z) = \Pr(Y_S = y \mid Z). \quad (4.20)$$

Dla wykluczających się oddziaływań t i c oraz wykorzystując fakt, że $\Pr(Y_c = y \mid S = t, Z) = 0$ oraz $\Pr(Y_t = y \mid S = c, Z) = 0$, na podstawie rachunku prawdopodobieństwa i prostej arytmetyki powyższa niezależność redukuje się do:

$$[\Pr(Y_t = y \mid S = t, Z) - \Pr(Y_c = y \mid S = c, Z)] \quad (4.21)$$

$$\Pr(S = c \mid Z) = 0.$$

W przypadku eksperymentu z randomizacją

$$\Pr(S = c \mid Z) = \Pr(S = c), \quad (4.22)$$

przy czym o $\Pr(S = c)$ można założyć, że przyjmuje w tym przypadku tylko wartości 0 lub 1.

Niezależność warunkowa redukuje się więc dalej do postaci

$$\Pr(Y_t = y \mid S = t, Z) = \Pr(Y_c = y \mid S = c, Z). \quad (4.23)$$

W przypadku jednak randomizacji S jest niezależne od Y_c oraz Y_t , a więc

$$\Pr(Y_t = y \mid Z) = \Pr(Y_c = y \mid Z). \quad (4.24)$$

Z tego z kolei wynika, że dla wszystkich wartości Z

$$E(Y_t \mid Z) = E(Y_c \mid Z). \quad (4.25)$$

Holland (1986b, 957) proponuje wprowadzenie definicji średniego efektu przyczynowego dla podpopulacji określonej przez warunek $Z = z$:

$$T(z) = E(Y_t - Y_c \mid Z = z). \quad (4.26)$$

Jeśli więc S nie jest przyczyną w sensie Grangera Y_S , to $T(z) = 0$ dla wszystkich wartości z . A więc w kontekście eksperymentu z randomizacją – brak przyczynowości w sensie Grangera implikuje zerowy średni efekt przyczynowy dla wszystkich podpopulacji określonych przez wartości zmiennej

Z. Natomiast jeśli t ma zerowy efekt dla wszystkich jednostek, wówczas w eksperymencie z randomizacją S nie będzie przyczyną w sensie Grangera dla Y_S względem żadnej zmiennej sprzed poddania oddziaływaniu.

Zasadniczą trudnością w odniesieniu do intuicyjnego pojęcia przyczyny jest – w przypadku Grangera – to, że dana zmienna może być zakwalifikowana do kategorii przyczyn tylko chwilowo, do momentu wprowadzenia innych zmiennych w analizie.

Mimo braku prac systematycznie podejmujących zagadnienie określenia związków modelu RCM z modelami równań strukturalnych SEM, należy za Hollandem podkreślić, że

zarówno modele przyczynowe [tj. SEM], jak i model Rubina, powstały, aby rozwiązać ten sam problem, jakim jest wnioskowanie przyczynowe w badaniach nieeksperymentalnych. (Holland 1986b, 958)

Jeśliby przyjąć, że dla każdego modelu SEM można przedstawić opisane w nim algebraicznie zależności przyczynowe w postaci zestawu równań strukturalnych graficznie, np. jako zorientowane krawędzie $X \rightarrow Y$, to na gruncie modelu RCM niektóre z takich diagramów nie posiadałyby interpretacji przyczynowej. Załóżmy, że A jest zmienną reprezentującą atrybut jednostki, a Y jest jej odpowiedzią. Wówczas otrzymalibyśmy mechanicznie pozbawiony interpretacji przyczynowej diagram $A \rightarrow Y$. Z drugiej jednak strony, jeśli S oznacza poddanie oddziaływaniu przyczyn, a Y_S jest zaobserwowaną odpowiedzią, to diagram $S \rightarrow Y_S$ miałby interpretację przyczynową.

Uogólniając (Holland 1986b, 958), można stwierdzić, że diagramy towarzyszące modelom SEM – dla dowolnej liczby zmiennych i z uwzględnieniem różnic w czasie pomiędzy oddziaływaniem poszczególnych zmiennych – można zrekonstruować w RCM. Warunkiem jest jednak to, aby diagramy te miały na gruncie RCM interpretację przyczynową. Próba dokonania takiej rekonstrukcji może więc uchronić zwolennika modeli SEM przed mechanicznym tworzeniem diagramów pozbawionych sensownej interpretacji przyczynowej.

4.3 Modele równań strukturalnych (SEM)

Poza klasyfikacją modeli przyczynowych zaproponowaną przez Coxa i Wermuth znalazło się jedno z najwcześniejszych podejść do modelowania

przyczynowości w badaniach nieeksperymentalnych, czyli modele równań strukturalnych.²⁷

Zasadnicze dla tego podejścia do modelowania jest wprowadzenie układu równań, które łącznie charakteryzują zależności między zmiennymi w nich występującymi. Każde z równań w modelu strukturalnym ma po lewej stronie zmienną (zależną, endogeniczną, odpowiedzi), dla której po prawej stronie równania umieszczono zmienne ją wyjaśniające oraz błąd losowy, który – przy jednej z interpretacji²⁸ – podsumowuje wpływ przyczynowy czynników nieuwzględnionych w modelu (i uznanych za zanedbywalne), o których zakłada się, że są wzajemnie niezależne.²⁹ Każda zmienna wyjaśniana występuje po lewej stronie tylko raz, natomiast może wystąpić jako zmienna wyjaśniająca we wszystkich pozostałych równaniach modelu strukturalnego.³⁰

Jak podkreśla Pearl (2000, 138), te „równania” z pozoru tylko przypominają symetryczność zwykłych równań algebraicznych, gdyż w przypadku równań występujących w modelach strukturalnych znak równości reprezentuje asymetryczność relacji przyczynowej.³¹ Zmienna po lewej stronie jest efektem oddziaływania zmiennych po prawej stronie. Przy takiej interpretacji te równania nie tylko reprezentują zaobserwowany rozkład, ale pozwalają także przewidywać, jaki efekt dla zmiennej po lewej stronie będą miały ma-

²⁷ Badając wykorzystanie technik analizy wielozmiennowej, S. Hershberger (2003) podkreśla, że jest to obecnie (dane za lata 1994-2001) najważniejsza z takich technik (w tych badaniach jako modele strukturalne ujęto także analizę ścieżkową, konfirmacyjną analizę czynnikową oraz struktury kowariancji) w porównaniu z innymi, jak: eksplanacyjna analiza czynnikowa, analiza grupowania, skalowanie wielowymiarowe, wielozmiennowa analiza wariancji MANOVA oraz analiza wyróżnikowa.

²⁸ Por. (Tintner 1938; Qin i Gilbert 2001). Do kwestii interpretacji błędów losowych powrócę poniżej, gdyż jest ona istotna dla dyskusji przyczynowej interpretacji zależności uwzględnianych w modelach strukturalnych.

²⁹ Jeżeli błędy losowe są zależne, wówczas można zinterpretować tę zależność jako sytuację, gdzie występuje ukryta wspólna przyczyna. Jeżeli bowiem błędy losowe reprezentują nieuwzględnione zależności, a jedna z nich sprawia, że te błędy losowe są wzajemnie zależne, to po uwzględnieniu odpowiedniej zmiennej otrzymalibyśmy model, w którym oba błędy losowe byłyby skutkami tej samej przyczyny, a jednocześnie względem niej byłyby od siebie warunkowo niezależne.

³⁰ Modele SEM przystępnie omawia J. Karpiński (1985, 35-47), a w ujęciu idealizacyjnej koncepcji nauki prezentuje je L. Nowak (1972, 118-28).

³¹ (Pearl 2000, 138; Lauritzen 2001, 74). Por. też (Mochon i Sloman 2004).

nipulacje zmiennymi występującymi po prawej stronie.³² Na tym polega strukturalny charakter parametrów występujących w równaniach – są niezmiennicze względem manipulacji dokonywanych na zmiennych.³³

Jak zauważają S. Greenland oraz B. Brumback, równania w modelach strukturalnych można potraktować jako „formuły pozwalające obliczyć potencjalne rezultaty dla różnych działań” (2002, 1034). Załóżmy, że X reprezentuje oddziaływanie, a U warunki, w których jest ono realizowane. Odpowiedź Y danej jednostki można więc przedstawić następująco:

$$Y = \beta_1 X + \beta_2 U + \varepsilon_Y. \quad (4.27)$$

Tak reprezentowaną zależność Greenland i Brumback odczytują jako wyraz potencjalnego rezultatu Y , jaki zostanie otrzymany przy np. zmianie wartości X i utrzymaniu U na tym samym poziomie. Wówczas z równania (4.27) można odczytać, że wartość Y zmieni się proporcjonalnie do iloczynu β_1 przez jednostki X , przy założeniu, że warunki U nie ulegną zmianie. Analogiczną interpretację można podać dla odwrotnej sytuacji, gdy zafiksowane jest X , a zmianie podlega U .

Jeżeli Równanie (4.27) jest tylko jednym z elementów modelu strukturalnego, to stwierdza ono także, że Y nie będzie zmieniać się przy zmianie wartości innych zmiennych, pod warunkiem, że X oraz U nie zostaną zmienione.

Na podstawie takiej interpretacji przyczynowej, podanej na poziomie poszczególnych jednostek, Greenland i Brumback odróżniają równania występujące w modelach strukturalnych od równań regresji:

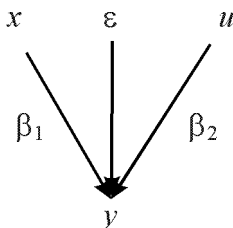
To właśnie te przyczynowe interpretacje na poziomie jednostkowym odróżniają równania strukturalne od zwykłych równań regresji (które reprezentują tylko *asocjacje* faktycznych rezultatów z faktycznymi wartościami współzmiennych dla przekroju jednostek). (2002, 1034)

Równaniom w modelu strukturalnym odpowiada model graficzny, w którym od wszystkich zmiennych występujących po prawej stronie danego równania wykreśla się strzałki skierowane w kierunku zmiennej wyjaśnianej,

³² Taki rodzaj manipulacji jest tylko jedną z możliwości, przy założeniu, że przy dokonywaniu interwencji na zmienne, manipuluje się tylko jedną z nich, a pozostałe są bez zmiany (Spirtes i Scheines 2005).

³³ Taki właśnie charakter miała wyjściowa interpretacja równań przez S. Wrighta. O powodach jej stopniowego zapoznawania piszę w Dodatku B.

występującej po jego lewej stronie. I tak, Równaniu (4.27) odpowiada graf przedstawiony na Rysunku 4.5.



Rysunek 4.5.

Ten sam graf (z pominięciem parametrów β_1 i β_2) reprezentuje także nieparametryczne równanie strukturalne o postaci:

$$Y = f_Y(X, U, \varepsilon_Y). \quad (4.28)$$

We wstępie do artykułu „Correlation and causation”, antycypującego to podejście, S. Wright podkreślił,³⁴ że

Idealna metoda nauki polega na wykazaniu bezpośredniego wpływu jednego warunku na drugi, w eksperymentach, w których wszystkie inne możliwe przyczyny zmienności zostały wyeliminowane. (1921, 557)

Jednak ze względu na to, że nie wszystkie z takich przyczyn można poddać kontroli w badaniach nieeksperymentalnych (tu: w naukach biologicznych):

często trzeba uwzględniać grupę charakterystyk lub warunków, które są skorelowane, z uwagi na złożoność wzajemnie oddziałujących na siebie, niekontrolowanych i niejasnych przyczyn. (1921, 557)

Zadanie, jakie stawia się temu sposobowi modelowania, Wright charakteryzuje krótko:

mierzenie bezpośredniego wpływu dla każdej oddzielnej ścieżki w takim [tj. złożonym i poddanym badaniu nieeksperymentalnemu] systemie i w ten sposób wskazanie stopnia, w którym zmienność danego skutku jest zdeterminowana przez każdą poszczególną przyczynę. (1921, 557)

³⁴ Zamysł Wrighta, pod nagłówkiem prawidłowości *ceteris paribus* w ekonomii, a szerzej w naukach społecznych oraz w filozofii, rozpowszechnił A. Marshall (1961, 36).

Ma on przede wszystkim doprowadzić do połączenia wiedzy przyczynowej, uzyskanej „a priori lub eksperymentalnie” i

Właśnie dlatego, że nie wymaga żadnych założeń co do natury związku [między zmiennymi] współczynnik korelacji można postrzegać jako fakt dotyczący opisu danej populacji, który można by jedynie kwestionować na podstawie niedokładności rachunkowej. Często jednak potrzeba wykorzystać jakąś metodę analizy, dzięki której można by połączyć wiedzę, jaką mamy o zależnościach przyczynowych, z wiedzą o stopniu zależności, jaką wyrażają współczynniki korelacji. (1921, 559)

W tym celu Wright wprowadził współczynniki ścieżek jako odzwierciedlenie siły związku przyczynowego między dwiema zmiennymi, niezależnie od wpływu na nie innych zmiennych. Współczynnik ścieżki dla zmiennych $X \rightarrow Y$ Wright zdefiniował – poprzez wariancję – jako stosunek standardowej wariancji odchylenia Y spowodowanej przez X do całkowitej standardowej wariancji odchylenia Y :

$$p_{X \rightarrow A} = \frac{\sigma_{X \cdot A}}{\sigma_X} \quad (4.29)$$

Standaryzacja pozwala obliczyć zmianę Y odpowiadającą dokonaniu zmiany w X o jedną jednostkę:

$$p_{X \rightarrow A} \frac{\sigma_X}{\sigma_A} = \frac{\sigma_{X \cdot A}}{\sigma_X} \quad (4.30)$$

Ogólnie, w modelu wielozmiennowym „korelacja między dwiema zmiennymi jest równa sumie iloczynów współczynników ścieżek dla każdej ze ścieżek, jakimi te zmienne są połączone” (1921, 568). Jedną z podkreślanych przez Wrighta zalet tego podejścia jest to, że współczynniki ścieżek można obliczyć, znając korelacje między zmiennymi (1921, 568).

Omówienie wypracowania koncepcji modeli strukturalnych w ekonometrii poprzedzę szczegółową dyskusją na temat Wrighta koncepcji analizy ścieżkowej. Charakterystyczne dla tego podejścia jest połączenie modelu algebraicznego równań równoważnych dla powiązanych przyczynowo zmiennych z jego reprezentacją graficzną. Na tę ostatnią składają się w głównej mierze skierowane w jedną stronę strzałki, reprezentujące bezpośrednie zależności przyczynowe – ich brak oznacza brak takiej zależności między zmiennymi. Jak podkreśla Wright (1960a, 190), część graficzna w

analizie ścieżkowej nie jest jedynie zilustrowaniem statystycznych faktów, lecz stanowi zasadniczy element analizy. Współczynniki ścieżek, identyfikowane na podstawie danych oraz równań modelu algebraicznego, są bowiem obliczane dla danego modelu graficznego i dlatego są wielkościami różnymi od zależności statystycznych dla poszczególnych zmiennych, które można odtworzyć na podstawie danych.³⁵

Każdy ze współczynników ścieżek ma

mierzyć ułamek standardowego odchylenia zmiennej zależnej, za który jest bezpośrednio odpowiedzialny dany czynnik. Chodzi o taki ułamek, który zostałby odkryty, gdyby ten czynnik zmieniał się w takim samym stopniu, jak w zaobserwowanych danych, podczas gdy wszystkie inne (łącznie z czynnikami resztowymi X_u) byłyby stałe. (Wright 1934, 162)

Współczynniki ścieżek pozwalają następnie na odniesienie korelacji między zmiennymi do określonych w modelu graficznym ich zależności funkcyjnych. To odniesienie dla dowolnych dwóch zmiennych V_0 i V_g Wright wyraża następującym wzorem (1934, 163):

$$r_{0g} = \sum P_{0i} r_{gi}, \quad (4.31)$$

gdzie P_{0i} jest współczynnikiem ścieżki od zmiennej V_i do zmiennej V_0 , a r_{XY} oznacza współczynnik korelacji dla zmiennych X oraz Y , czyli we współcześnie używanej notacji $\rho_{XY} = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$.

Zgodnie ze wzorem (4.31), można rozłożyć korelację między zmiennymi X oraz Y na „wkład wszystkich ścieżek w danym diagramie”, które prowadzą od jednej z nich do drugiej (Wright 1934, 163).

Zwróćmy w tym miejscu uwagę na podkreślaną przez Wrighta – szczególnie wyraźnie w (1960, 444-45) – istotną cechę analizy ścieżkowej, która wydaje się zasadniczo różnić ją od współcześnie proponowanych modeli graficznych przyczynowości, a więc modeli bayesowskich w podejściu Pearl'a oraz Glymour'a. Wright zwraca uwagę na to, że współczynniki ścieżek nie są absolutną własnością analizowanych zmiennych, gdyż są wtórne wobec przyjętych między nimi zależności funkcyjnych, jakie wyraża model graficzny. Ten model graficzny, podkreślmy za Wrightem, nie jest wyrowadzony z danych, lecz przyjęty niezależnie. Analiza korelacji zgodnie z

³⁵ To także wydaje się elementem modelowania Wrighta, na który nie zwracają uwagi Pearl czy Glymour, formułując algorytm odtwarzania modelu graficznego na podstawie danych (korelacji cząstkowych).

modelem graficznym zależności jest zatem wtórna wobec tego modelu. Zgodnie z rozpowszechnionym dziś podejściem Pearla, a zwłaszcza Glymoura i współpracowników, jest przeciwnie – modele graficzne są wyprowadzane z danych poprzez analizę korelacji cząstkowych między zmiennymi. Do tej kwestii powrócę w Rozdziale 5.

Wzór Wrighta (4.31) jest rekursyjny, a więc pozwala na dalszą analizę występującego w nim terminu „ r_{gi} ” zgodnie z podaną zależnością. Wright wprowadza więc pojęcie – wyrażane współcześnie terminem „kolider” – zablokowania ścieżki przez zmienną, w kierunku której skierowane są dwie przeciwstawnie skierowane strzałki:

Przy wyszukiwaniu łączących ścieżek oczywiste jest, że można iść w kierunku przeciwnym do skierowania strzałek, a następnie w kierunku zgodnym z ich skierowaniem oraz że można przejść bezpośrednio od jednej zmiennej do drugiej (być może za pośrednictwem zmiennych zakłócających), ale *nigdy nie można iść zgodnie, a potem niezgodnie ze skierowaniem strzałek.* (Wright 1934, 163; podkreślenie – P.K.)

Uzasadnienie, dlaczego koliderzy – czy w terminologii Pearla „struktury V” – odgrywają taką właśnie rolę, podam pełniej za autorami graficznych bayesowskich modeli przyczynowości w Punkcie 4.4. W tym miejscu przywołam tylko zdawkowe wyjaśnienie, jakie daje Wright: „To, że dwa czynniki wpływają na tę samą zmienną zależną, nie tworzy między nimi korelacji” (1934, 163).

W modelu graficznym dopuszczalne są także dwukierunkowe strzałki. Mają one wyrażać zależność między nieuwarunkowanymi czynnikami przyczynowymi (zmienne egzogeniczne) dla różnych efektów przyczynowych. A więc dla różnych czynników determinujących ten sam skutek (tę samą zmienną endogeniczną) można w analizie ścieżkowej zawsze przyjąć, że są statystycznie niezależne. Takiego jednak założenia, bez podania uzasadniającej je – i niezależnej od specyfikacji modelu – argumentacji, nie można przyjąć w stosunku do zmiennych determinujących różne skutki (zmiennych egzogenicznych).

Wright, dyskutując różnice między analizą ścieżkową a regresją oraz wskazując ich komplementarność, podsumowuje zagadnienie estymacji współczynników w analizie ścieżkowej następująco:

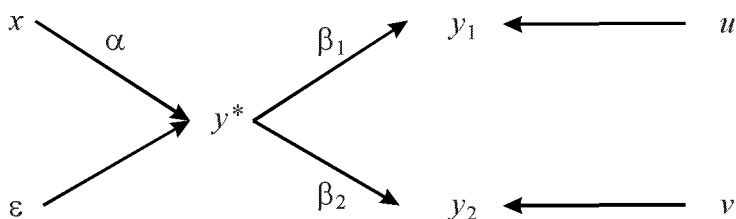
korelacja między dowolnymi dwiema zmiennymi we właściwie skonstruowanym diagramie relacji jest równa sumie wkładu poszczególnych ścieżek, którymi można przejść od jednej do drugiej w tym diagramie bez zawracania na wybranej strzałce i bez dwukrotnego uwzględniania danej zmiennej na tej samej ścieżce. Współczynnik dla całej ścieżki łączącej dwie zmienne, a tym sa-

mym mierzący wkład tej ścieżki do korelacji, znany jest jako *złożony współczynnik ścieżki*. Jego wartość jest iloczynem wartości poszczególnych współczynników elementarnych na złożonej ścieżce, łączącej dwie zmienne. (Wright 1960a, 193)

Jako ilustrację przywołam jeden z przykładów, jakie podaje Wright (1960a, Rysunek 5, 199), w postaci zmodernizowanej przez Goldbergera (1972, 987). Model algebraiczny, gdzie x jest zmienną egzogeniczną, y_1 oraz y_2 są endogeniczne, y^* jest zmienną nieobserwowalną (o zestandaryzowanej wariancji $\sigma_{**} = 1$), a zakłócenia ε , u oraz v są niezależne, przedstawia się następująco:

$$y^* = \alpha x + \varepsilon, \quad y_1 = \beta_1 y^* + u, \quad y_2 = \beta_2 y^* + v. \quad (4.32)$$

Graficzny model ścieżkowy w tym przypadku jest następujący:



Rysunek 4.6.

Brak dwukierunkowych strzałek na tym diagramie reprezentuje przyjęte założenie o niezależności między zmienną egzogeniczną a zmiennymi zakłócającymi. Zgodnie z wyżej podaną przez Wrighta wskazówką, z tego diagramu można odczytać zależności dotyczące wartości parametrów i wartości rozkładu dla populacji (wariancje i kowariancje dla populacji):

$$\sigma_{11} = \beta_1^2 + \sigma_{uu}, \quad \sigma_{12} = \beta_1\beta_2, \quad \sigma_{1x} = \alpha\beta_1\sigma_{xx}, \quad (4.33)$$

$$\sigma_{22} = \beta_2^2 + \sigma_{vv}, \quad 1 = \alpha^2\sigma_{xx} + \sigma_{\varepsilon\varepsilon}, \quad \sigma_{2x} = \alpha\beta_2\sigma_{xx}.$$

Występujące w tych równaniach parametry (α , β_1 , β_2 , σ_{uu} , σ_{vv} , $\sigma_{\varepsilon\varepsilon}$) mogą być estymowane po wprowadzeniu do równań w (4.33) wartości zaobserwowanych w próbie.

W zastosowaniach analizy ścieżkowej Wright, bazując na przedstawionych tu podstawowych elementach tej analizy, skonstruował złożone modele, służące do badania takich zjawisk, jak rozmiary kości królików, kolor skóry i waga urodzeniowa świnek morskich, rozprzestrzenianie się roślin, ludzka inteligencja a czynniki wrodzone i środowiskowe, struktura genetyczna populacji itp. Jedną z zasadniczych zalet tego podejścia jest wprowadzenie estymacji dla zmiennych nieobserwowalnych, jak y^* w omówionym przykładzie, które występują w modelach.

Jednym z ograniczeń podejścia Wrighta, na które zwraca uwagę Goldberger, jest to, że w przypadku nadidentyfikowalności (gdy jest więcej równań estymujących niż parametrów do estymowania)³⁶ podejście Wrighta nie daje precyzyjnych wskazówek, jak osiągnąć najlepszą estymację. Goldberger charakteryzuje postępowanie Wrighta w takich przypadkach jako *ad hoc* (uśrednianie kilku równań lub nieformalne sprawdzanie, czy różne estymacje dają bliskie wartości) i ponadto, zarzuca mu, że „rzadko podaje on ścisły model statystyczny i często niejasno posługuje się odróżnieniem między wartościami z próby i populacji” (1972, 988).

Aplikację analizy ścieżkowej do badań socjologicznych zaproponował w 1966 r. D. Duncan.³⁷ Zastosował ją w modelach społecznych uwarunkowań ilorazu inteligencji, rozwoju miast i gęstości zaludnienia, rankingów prestiżu zawodowego oraz międzypokoleniowej mobilności. Jak podkreśla Goldberger, w tych, jak i późniejszych zastosowaniach metody Wrighta w socjologii, powielane są wspomniane wcześniej ograniczenia tego podejścia.

Warto także wspomnieć, że modele strukturalne występowały w innych jeszcze formach, z których szeroko znana jest analiza czynnikowa stosowana w badaniach psychologicznych do lat 1930. XX wieku.³⁸ Typowym przykładem takiego badania jest wyjaśnianie wyniku uzyskanego przez studenta w

³⁶ To zjawisko Wright wykorzystywał do testowania podanych modeli (Goldberger 1972, 989).

³⁷ H. Blalock (1971) dokonał przeglądu tekstów, gdzie stosuje się analizę ścieżkową w innych dyscyplinach. Por. też (Christ 1960; Hildreth 1960; Jahoda i in. 1965; Duncan 1975; Freedman 1983; Holland 1988).

³⁸ Glymour szczegółowo dyskutuje osiem założeń, na jakich oparte jest poprawne stosowanie analizy czynnikowej. Podważa trzy z nich: 1) wszystkie korelacje są wyłącznie pochodną niezmiernych wspólnych przyczyn; 2) funkcja, określająca zależność między zaobserwowanymi odpowiedziami a ukrytymi czynnikami, jest taka sama dla każdej osoby; oraz 3) własności osób, które generują odpowiedzi na pytania w testach psychometrycznych, są stałe dla każdej osoby (2001, 187).

teście przez niewielką liczbę nieobserwowalnych zdolności umysłowych (Goldberger 1972, 990-93).

Rozwój technik charakterystycznych dla modeli strukturalnych to zasługa ekonomistów skupionych w Komisji Cowlesa, mającej na celu zintegrowanie ekonomii i matematyki.³⁹ W latach 1940. do tej komisji dołączył T. Haavelmo, któremu przypisuje się udane zintegrowanie wcześniejszych idei i wypracowanie nowoczesnej idei modelowania strukturalnego.

Haavelmo, proponując własną koncepcję modelowania w ekonomii,⁴⁰ przeciwstawił się dotychczasowej:

Dotychczas powszechnie stosowana procedura polega na rozpoczęciu od skonstruowania teorii ekonomicznej, określającej *dokładne* zależności funkcyjne, a następnie porównaniu tej teorii z pewnymi faktycznie dokonаныmi pomiarami, by na końcu „ocenić”, czy korespondencja między nimi jest „dobra”, czy „zła”. (1944, iii)

W takim podejściu do modelowania Haavelmo nie dostrzega możliwości zasadnego stosowania narzędzi wnioskowania statystycznego, gdyż te wymagają, by modele były nie deterministyczne, lecz stochastyczne.⁴¹ Powodem dotychczasowego odrzucania takich modeli zjawisk ekonomicznych było, zdaniem Haavelma, to, że naturę tych zjawisk uważano za niestochastyczną, a po drugie, jeśli nawet by dopuszczano stochastyczność, to żądano, aby podstawą modelu probabilistycznego była niezależność obserwacji. To ostatnie jest warunkiem bardzo mocnym – i rzadko możliwym do spełnienia – gdyż oznacza, że kolejne obserwacje zjawisk ekonomicznych miałyby pochodzić z tej samej „populacji”. W odpowiedzi na te zarzuty Haavelmo odpowiada, dając zwięzły zarys swojej koncepcji:

³⁹ Tę komisję w 1932 r. ustanowił Alfred Cowles w Colorado Springs, skąd w 1939 r. przeniesiona została do Chicago, gdzie funkcjonowała do 1952 r. Jej zasadniczy wkład to wypracowanie podstawowej aparatury pojęciowej dla ekonometrii. Wśród osób związanych z tą komisją znalazło się wielu późniejszych laureatów Nagrody Nobla: Tjalling Koopmans, Kenneth Arrow, Gerard Debreu, James Tobin, Franco Modigliani, Herbert Simon, Lawrence Klein, Trygve Haavelmo i Harry Markowitz. Szczegóły historyczne prac tej komisji szerzej omówione są na jej oficjalnej stronie w Internecie: <http://cowles.econ.yale.edu>.

⁴⁰ Haavelmo podkreśla, że spośród różnych czynności badawczych, jak konstruowanie modeli, testowanie teorii, estymacja oraz predykcja, najważniejsza jest dla ekonomistów „konstrukcja racjonalnych modeli, które przyczynią się w jakiś sposób do naszego rozumienia rzeczywistego życia ekonomicznego” (1944, 11). Stąd też przede wszystkim tej właśnie czynności poświęca swoją pracę.

⁴¹ Por. m.in. (Hurwicz 1994).

Nie jest konieczne, aby obserwacje były niezależne i aby wszystkie były zgodne z tym samym jednowymiarowym prawem prawdopodobieństwa. Wystarczy bowiem założyć, że *cały zbiór* obserwacji, np. n , może być uznany za *jedną* obserwację n zmiennych (lub „punkt w próbie”), zgodnych z n -wymiarowym prawem prawdopodobieństwa *łączniego*, którego „istnienie” może być czysto hipotetyczne. (1944, iii)

Takie „łączone prawo prawdopodobieństwa” pozwala na wnioskowanie statystyczne i testowanie modelu dzięki „jednemu punktowi z próby”. Bardzo pomocnych w tym narzędzi dostarcza wypracowana przez Neymana i Pearsona teoria wnioskowań statystycznych.⁴²

Haavelmo podkreśla także, iż w dotychczasowej praktyce ekonomistów, nawet tych przeciwnych modelowaniu stochastycznemu, można znaleźć elementy zastosowania pojęcia prawdopodobieństwa oraz zmiennych losowych. Formułując bowiem swoje teorie, nie oczekują oni na uzyskanie pełnej zgodności z danymi, klasyfikując ją w kategoriach „dopuszczalnej” i „praktycznie niemożliwej”. Te dwie kategorie, zdaniem Haavelma, można zastąpić bardziej ugruntowanymi teoretycznie kategoriami probabilistycznymi: „prawdopodobieństwo bliskie 1” oraz „prawdopodobieństwo bliskie 0” (1944, iv).

Argumentem, który podaje Haavelmo, za taką rekonceptualizacją dotychczasowego podejścia do modelowania zjawisk ekonomicznych, jest to, że otrzymujemy w ten sposób twierdzenia, które można testować. Wykorzystując więc zaproponowaną rekonceptualizację, otrzymamy zdania dotyczące faktycznych zjawisk, np. A , o postaci „prawdopodobieństwo, że A wystąpi, jest bliskie 1”. Takie zdania są testowalne na gruncie teorii wnioskowania statystycznego opracowanej przez Neymana-Pearsona.

Ostatecznie jednak taka rekonceptualizacja jest uzasadniona przez Haavelma własnościami logicznymi proponowanych modeli:

Jakakolwiek byłaby teoria, nie może ona pozostać prawdziwa w odniesieniu do pewnego zbioru faktów, jeśli kiedykolwiek *implikuje* fałszywe twierdzenie dotyczące tych faktów. Uważamy więc za praktycznie niemożliwe utrzymywanie jakiegokolwiek teorii, która *implikuje* nietyrwialne stwierdzenie dotyczące pewnych faktów, ponieważ wcześniej czy później fakty, na ogół, zaprzeczają temu twierdzeniu. Wobec tego [...] przyjmujemy szczególny rodzaj

⁴² Heckman podkreśla, że synteza programu Komisji Cowlesa z teorią wnioskowań statystycznych Neymana-Pearsona wydaje się nieunikniona, gdyż ówczesnie w zasadzie była ona jedyną szczegółowo opracowaną taką teorią (2000, 86). Jej konsekwencją było jednak odseparowanie konstrukcji modelu od pozyskiwania danych, co jest niezgodne z „interakcyjną naturą empirycznego procesu uczenia się” (2000, 87).

modelu, a mianowicie dopuszczającego stwierdzenia, które nie są *implikacjami*, a mają jedynie pewną szansę bycia prawdziwymi. To doprowadzi nas do probabilistycznego formułowania teorii, które mają znaleźć zastosowanie. (1944, 2)

Podsumowując, Haavelmo podkreśla, że formułowanie modeli ekonomicznych za pomocą hipotez statystycznych, dotyczących rozkładu prawdopodobieństwa dla badanych zjawisk, jest koniecznością, jeśli chce się wykorzystać narzędzia wnioskowania statystycznego, a jednocześnie nie prowadzi to do zmniejszenia ogólności proponowanych stwierdzeń (1944, iv).

Przywołajmy wprowadzone Punkcie 2.2 odróżnienie prawdopodobieństwa *ontycznego* od *epistemicznego*. Haavelmo, podając uzasadnienie stochastycznego formułowania modeli ekonomicznych, wykorzystuje ten drugi rodzaj prawdopodobieństwa. W przytoczonej tu argumentacji nie charakteryzuje on zjawisk ekonomicznych (a dokładniej zależności przyczynowych, jakie zachodzą między nimi) jako wsobnie probabilistycznych.⁴³

Zasadniczą racją, jaką podaje, jest własność twierdzeń wyprowadzanych z takiego modelu, które – aby można było je uzgodnić z danymi doświadczenia – muszą mieć postać, zgodnie z którą „mają jedynie szansę bycia prawdziwymi lub fałszywymi”. Dlatego też modele, z których są wyprowadzane, muszą być wsobnie probabilistyczne, a nie deterministyczne. Taki probabilistyczny charakter modeli jest w tym przypadku motywowany kryteriami prawdziwości (a nie ontycznymi własnościami modelowanych zjawisk i zachodzących między nimi zależności), z czego wynika, jak sądzę, jego *epistemiczny* charakter. Do bardziej szczegółowej dyskusji tej kwestii powrócę poniżej, po zarysie teorii modelowania, jaką podał Haavelmo.

Modele strukturalne, w ujęciu Haavelma, służą określeniu zależności zachodzącej dla grupy zmiennych reprezentujących zjawiska ekonomiczne. Zależności występujące w tych modelach dzielą się na trzy zasadnicze grupy: definicje (np. wydatek całkowity = cena pomnożona przez zakupioną ilość danego produktu), relacje techniczne (np. przyrodnicze czy instytucjonalne ograniczenia uwzględnianie w planowaniu ekonomicznym) oraz za-

⁴³ Wprowadzenie przez Haavelma prawdopodobieństwa w sensie epistemicznym, a nie ontycznym, jest spójne z jego filozoficznie fundamentalnymi poglądami dotyczącymi praw przyrody: „nie można opisać czegoś takiego, jak prawo przyrody, bez odwołania się do określonych *zasad analizy*” (Haavelmo 1944, 12). Takie przekonanie jest wyrazem wyrafinowanego połączenia pozytywistycznego z empirystycznym podejściem w nauce, wypierając prymitywny indukcjonizm: zbieranie faktów, które następnie miałyby być apriorycznie kategoryzowane za pomocą aparatury analitycznej (matematycznej).

leżności charakteryzujące działania ekonomiczne (np. opis decyzji jednostek lub grup w odniesieniu do produkcji czy konsumpcji).

Oprócz zmiennych, w tych modelach występują także parametry, które są wielkościami wprowadzanymi w procesie analizy zjawisk reprezentowanych przez zmienne. Haavelmo (1944, 5-7) szczegółowo omawia różnicę między zmiennymi występującymi w modelu „wyjaśniającym” zjawiska ekonomiczne a zmiennymi występującymi w modelu faktycznie stosowanym w badaniu, gdzie trzeba uwzględnić problemy z niedokładnością klasyfikacji zjawisk oraz przybliżonością ich pomiaru. Zaznacza przy tym, że kwestie związane z proceduralnością, tak jak scharakteryzowałem ją w drugim rozdziale tej książki, zostają poza obszarem modelowania i dotyczą wyłącznie kwestii aplikacji modeli do badanej rzeczywistości:

powinno się bardzo uważnie przestudiować rozważane [w badaniu] faktyczne szeregi i warunki, w jakich powstały, przed utożsamieniem ich ze zmiennymi w jakimś modelu teoretycznym. (1944, 7)

Haavelmo zauważa jednak, że „Taki model teoretyczny, rozważany sam w sobie, nie ma żadnego znaczenia praktycznego” (1944, 8). W takim podejściu do modelowania konieczne jest uwzględnienie – zewnętrznego w stosunku do modelu – odniesienia do rzeczywistych zjawisk:

Model nabiera znaczenia ekonomicznego tylko wtedy, gdy wybrano lub opisano odpowiadający mu system wielkości lub obiektów w rzeczywistym życiu ekonomicznym po to, aby *utożsamić je z tymi* w modelu. A więc model będzie miał ekonomiczne znaczenie tylko wtedy, gdy będzie skojarzony z planem faktycznych eksperymentów, który opisuje – oraz wskazuje, jak mierzyć – system „prawdziwych” zmiennych (lub przedmiotów) [...], które mają zostać utożsamione z odpowiednimi zmiennymi w teorii. (1944, 8)

Zwróćmy uwagę na podkreślony przez Haavelma w zacytowanym tekście fakt, że zagadnienie planowania badań („eksperymentu”), które ma prowadzić do uzyskania informacji o badanych zjawiskach ekonomicznych, samo nie jest elementem modelu zależności wiążących te zjawiska. A zatem, interpretacja przyczynowa planu badań jest wprowadzana niezależnie od tego, jakie dane uzyskuje się w interakcji z reprezentowanymi w takim modelu zjawiskami. W tym sensie więc dla badań nieeksperymentalnych propozycja Haavelma jest niezgodna z zaproponowanym tu podejściem proceduralnym. W konsekwencji bowiem „taki model staje się *hipotezą a priori* o rzeczywistych zjawiskach”, a wszystkie obserwowalne wartości mają leżeć w granicach wyznaczonych teoretycznie (Haavelmo 1944, 9). Do tej kwestii powrócę w kolejnym rozdziale.

Podjmując problem istnienia niezmienniczych zależności między zmiennymi reprezentującymi zjawiska ekonomiczne, Haavelmo odnosi się do trzech kwestii (1944, 17): odwracalności relacji ekonomicznych, prostoty w formułowaniu praw ekonomicznych oraz do autonomii relacji ekonomicznych.⁴⁴

Zagadnienie przyczynowości w rozważaniach Haavelma dyskutowane jest w odniesieniu do kwestii prostoty w formułowaniu praw ekonomicznych. Jeśli Y oznacza rezultat decyzji ekonomicznej podjętej przez jednostkę lub firmę (np. roczna konsumpcja pewnego produktu przez określoną grupę jednostek), to analizę ma uprościć założenie, że na tę zmienną oddziałują określone czynniki przyczynowe. Nie chodzi tu jednak o „metafizyczny sens” przyczynowości,⁴⁵ lecz o to, że „jednostki, firmy itp. są ograniczone w swoim planowaniu i decyzjach przez zbiór warunków, które są *danymi* w procesie adaptacji” (1944, 22). W ramach więc upraszczającego założenia „przyczynowego” przyjmuje się, że dla każdego zbioru warunków (zmienne niezależne) podmioty ekonomiczne mają określony system preferencji, które przesądzają, co w tych warunkach jest najlepszą (i jedyną) dla nich decyzją (zmienne zależne).

Haavelmo podkreśla dalej, że posługiwanie się takim założeniem przyczynowym nie wyklucza modelowania wzajemnych zależności między zjawiskami (np. warunkami a decyzjami podmiotów). Dla wyświetlenia tej wzajemności należy rozłożyć modelowaną sytuację na takie przypadki, gdy te zależności są jednokierunkowe, choć w każdym z tych przypadków mogą być zorientowane przeciwnie. W różnych kontekstach te same zmienne mogą więc pełnić różne role: zmiennych zależnych w jednym, a niezależnych w innym kontekście.⁴⁶

Przy takiej interpretacji zmiennych zależnych i niezależnych zagadnienie prostoty modelowania sprowadza się więc do tego, żeby

⁴⁴ Jest to jedna z kluczowych kwestii dla zaproponowanego przez Pearla „uogólnienia i abstrakcji” modeli strukturalnych, jaką są bayesowskie modele zależności przyczynowych. Dyskutowaną przez Haavelma własność zależności przyczynowych Pearl określa jako ich autonomiczność (Pearl 2000, 22-23).

⁴⁵ W dalszej części tekstu Haavelmo jednak podważa poniekąd tę deklarację, pisząc o modelach jako uproszczonej rekonstrukcji „mechanizmów, o których sądzimy, że są ukryte pod zjawiskami, jakie obserwujemy w świecie rzeczywistym” (Haavelmo 1944, 26).

⁴⁶ Pearl, podając przykład modelu przyczynowego dla zjawisk popytu i podaży, interpretuje te dwa modele jako reprezentujące różne mechanizmy przyczynowe dla tych samych zjawisk (Pearl 2000, 27-29).

dla każdej zmiennej Y , która ma zostać „wyjaśniona”, istniało stosunkowo niewiele wyjaśniających czynników, których zmienność praktycznie determinuje zmienność Y . (1944, 23)

Spełnienia tego postulatu nie można zagwarantować apriorycznie, jednak w celu ułatwienia analizy Haavelmo wprowadza definicje wpływu potencjalnego i faktycznego. Ten pierwszy obliczany jest dla jednej zmiennej niezależnej na podstawie podanej w modelu zależności danego zjawiska od czynników przyczynowych i przy założeniu, że wszystkie pozostałe zmienne niezależne się nie zmieniają. Wpływ faktyczny natomiast obliczany jest niezależnie od podanej w modelu zależności, a więc z uwzględnieniem podanych przez „zewnątrzną w stosunku do modelu zasadę”, która określa różne rodzaje zmiennych, jakie mogą wpływać na dane zjawisko, oprócz tych, które występują w modelu. Następnie określona jest miara sumująca ten wpływ w różnych uwzględnionych przypadkach, przy założeniu, że zmienia się jedynie interesująca nas zmienna niezależna, a pozostałe mają stałą wartość (Haavelmo 1944, 23-24).

Jako ogólną zasadę Haavelmo proponuje więc pominięcie w specyfikacji modelu wszystkich tych zmiennych, których łączny faktyczny wpływ na zmienną zależną jest „bardzo mały” w porównaniu z faktycznym wpływem zmiennych uwzględnionych w modelu. Taka sytuacja może zaistnieć w jednym z dwóch przypadków: (1) czynniki nieuwzględnione w modelu mają mały wpływ potencjalny bądź (2) mają duży wpływ potencjalny, ale zmieniają się nieznacznie lub rzadko w porównaniu z czynnikami uwzględnionymi w modelu.

Niejako odwrotnością problemu prostoty jest problem określony przez Haavelma jako autonomiczność relacji ekonomicznych.⁴⁷ Dla ilustracji tego zagadnienia, przywołam za Haavelmem analogię z jazdą samochodem (1944, 27-28). Na pytanie o to, jaka jest zależność między naciskaniem pedału gazu a prędkością jazdy, można dać różne odpowiedzi. Po wielu próbach jazdy na płaskiej i suchej trasie można sformułować dość dokładną zależność funkcyjną między siłą, z jaką naciska się pedał gazu, a maksymalną prędkością jazdy. Taka wiedza powinna być wystarczająca do prowadzenia tego samochodu na tej trasie z wybraną prędkością. Taka wiedza nie pozwala jednak zrozumieć, jak działa samochód. Po pierwsze, nic nie mówi o sposobie funkcjonowania mechanizmu, jaki jest w silniku tego samochodu, a po drugie – określona zależność funkcyjna między naciskaniem pedału

⁴⁷ Termin „autonomiczność” pochodzi z pracy R. Frischa „Autonomy of economic relations” przygotowanej w 1938 r. dla Ligi Narodów.

gazu a prędkością zmieni się wraz ze zmianą funkcjonowania jednej z części silnika lub jego usterką.⁴⁸ Tę zależność funkcyjną Haavelmo określa jako nieautonomiczną, gdyż „jej istnienie zależy od jednoczesnego spełnienia wielkiej liczby innych zależności, wśród których są zależności o nietrwalej naturze” (1944, 28).

Przeciwieństwem nieautonomicznej zależności funkcyjnej są ogólne prawa termodynamiki, tarcia. Opisują one funkcjonowanie pewnych części w mechanizmie samochodu niezależnie od tego, jak funkcjonują inne. I dlatego Haavelmo obdarza je mianem autonomicznych.

Ta analogia ma, zdaniem Haavelma, bezpośrednie odniesienie do „mechanizmów życia ekonomicznego”:

Teoria ekonomiczna bazuje na założeniu, że decyzje jednostek o produkcji i konsumpcji można opisać za pomocą określonych relacji behawiorystycznych, a ponadto, że istnieją pewne techniczne i instytucjonalne ograniczenia wolności wyboru (jak np. techniczne funkcje produkcji, ograniczenia prawne itp.).

Określony system takich relacji definiuje jedną teoretyczną *strukturę* w ekonomii, a mianowicie, definiuje teoretyczny *zbiór* możliwych jednoczesnych zbiorów wartości lub zbiorów szeregów czasowych dla zmiennych ekonomicznych. (1944, 28)

Różnica między konstrukcją silnika samochodu a „mechanizmem życia ekonomicznego” dotyczy tego, że w tym ostatnim przypadku teoria ekonomiczna musi niekiedy uwzględnić alternatywne struktury zależności, które w przybliżony sposób korespondują z rzeczywistością ekonomiczną.

Określenie autonomiczności jednej relacji względem innej nie jest kwestią analizy matematycznej czy logicznej. Haavelmo podkreśla, że

to jest problem faktycznej *wiedzy o rzeczywistych zjawiskach* i przyjmowania realistycznych założeń w odniesieniu do nich. Próbując ustanowić relacje o wysokim stopniu autonomii, bierzemy pod uwagę różne *zmiany* w strukturze ekonomicznej, które mogą zmienić nasze stosunki. Staramy się wydobyć takie relacje, od których można faktycznie oczekiwać, że będą miały wysoki stopień niezmienniczości w stosunku do pewnych zmian w strukturze, które są „rozsądne”. (1944, 29)

Zwróćmy uwagę, że w ujęciu Haavelma – w przeciwieństwie do współczesnych prac, zwłaszcza Pearla – autonomiczność relacji jest cechą

⁴⁸ I. Niiniluoto podaje parafrazę tego przykładu, którą przytaczam w Zakończeniu jako element jego argumentacji za realizmem naukowym; por. (Niiniluoto 1999, 197-98).

względna, a „konstrukcja systemów relacji autonomicznych jest kwestią intuicji i wiedzy faktualnej, jest sztuką” (1944, 29).

Autonomiczności relacji, jak podkreśla Haavelmo, nie można wywnioskować z obserwacji, która daje wiedzę tylko o trwałości danej zależności w odniesieniu do faktycznych obserwacji, a tym samym zależy od tego, jakie zmiany faktycznie wystąpiły. Autonomiczność relacji natomiast „odnosi się do klasy *hipotetycznych zmian* w strukturze, w stosunku do których ta relacja *byłaby niezmienna*” (1944, 29).

Dyskusję na temat autonomiczności zależności ekonomicznych Haavelmo konkluduje uwagą, że „fundamentalne zależności” są niezmiennicze względem „znacznie szerszej klasy zmian niż te, które ujawniają się nam w naturalnym biegu zdarzeń” (1944, 38). Stąd też występuje problem estymacji takich fundamentalnych zależności na podstawie bieżących obserwacji.

W odniesieniu do prób wyprowadzenia zależności przyczynowych z danych można przywołać następujące stwierdzenie Haavelma:

W przypadku zmian, które obserwujemy, możliwe jest ustanowienie nieskończonej ilości zależności przez po prostu łączenie dwóch lub więcej fundamentalnych zależności na różne sposoby. [...] To jest oczywiście jeszcze jeden aspekt ogólnej reguły [...]: reguły, że każdej teorii powinno towarzyszyć planowanie eksperymentów. (1944, 39)

Z perspektywy proceduralnej można uzupełnić ten komentarz Haavelma o uwagę, że planowanie badań jest elementem nieodzownym w odkrywaniu zależności przyczynowych.

Zasadnicze dokonanie Haavelma w teorii modeli strukturalnych to zintegrowanie w niej pojęć probabilistycznych, a przez to także umożliwienie zastosowania opracowanej przez Neymana i Pearsona teorii wnioskowań statystycznych do tych modeli, przede wszystkim testowania, estymacji i predykcji (1944, 60-113). Swoją koncepcję Haavelmo szczegółowo prezentuje w rozdziale trzecim *The probability approach in econometrics*, który zatytułował „Stochastical schemes as a basis for econometrics”.

Chociaż w tym miejscu Haavelmo nie zaznacza tego w sposób jawny, punktem wyjścia w modelowaniu jest dla niego sformułowanie teoretycznych funkcyjnych zależności dla zmiennych „teoretycznych”: „jeśli rozpoczynamy od schematu teoretycznego, musimy – z uwagi na zastosowania – dodać pewne elementy stochastyczne, aby wypełnić lukę między teorią a faktami” (1944, 52).

Dopiero kolejny krok polega na utożsamieniu ich ze zmiennymi reprezentującymi zaobserwowane (zmierzone) zjawiska. To utożsamienie przebiega z kolei następująco:

apriorycznie możemy powiedzieć, że teorii [...], że nie wolno *wykluczyć z grona możliwych* którekolwiek z już zaobserwowanych systemów wartości „prawdziwych” zmiennych, ani żadnego, który jest praktycznie dopuszczalny w przyszłości. (1944, 40)

Bez dodatkowych wymogów w stosunku do takiego utożsamienia otrzymalibyśmy modele „za szerokie”, które „nie miałyby praktycznego zastosowania”. Dlatego

potrzeba nam teorii, które – bez wprowadzania bezpośrednich logicznych sprzeczności – stwierdzą, że obserwacje *z reguły* zgrupują się w ograniczonym podzbiórze zbioru wszystkich pojmowalnych obserwacji, podczas gdy z tą teorią nadal zgodne będzie to, że jakaś obserwacja „od czasu do czasu” wypadnie poza ten podzbiór. (1944, 40)

Haavelmo w tym miejscu, podobnie jak wielu mu współczesnych ekonomistów, zwłaszcza R. Frisch i T. Koopmans (Hendry i Morgan 1995, 33-45), za wzór modelowania przyjmuje nauki ścisłe, a przede wszystkim astronomię,⁴⁹ traktując odchylenia od zadanego apriorycznie modelu teoretycznego jako aberracje.

Haavelmo jednoznacznie sygnalizuje więc, że nie wprowadza w modelach pojęcia prawdopodobieństwa w sensie ontycznym, jako własności samych reprezentowanych w modelu zjawisk,⁵⁰ lecz w sensie epistemicznym, jako wyraz naszej niewiedzy o wszystkich zależnościach w badanej rzeczywistości:

Kwestia nie dotyczy tego, czy prawdopodobieństwa *istnieją*, czy nie, lecz tego, byśmy postępując *jak gdyby* istniały, byli w stanie podać takie stwierdzenia o rzeczywistych zjawiskach, które są „poprawne dla celów praktycznych”. (1944, 43)

⁴⁹ Krytykę stosowania narzędzi modelowania i analizy danych obserwacyjnych przenoszonych bezpośrednio z nauk przyrodniczych na grunt nauk społecznych podał Freedman (1987a; 1987b); por. też (Achen 1987; Bentler 1987; Cliff 1987; Fox 1987; Hope 1987; Karlin 1987; Muthen 1987; Rogosa 1987; Rothenberg 1987; Seneta 1987; Wold 1987). Do tej dyskusji powróć w kolejnym rozdziale.

⁵⁰ Pojęcie prawdopodobieństwa interpretuje on częstościowo, jako stosunek zaobserwowanej liczby interesujących przypadków do wszystkich zaobserwowanych w danej grupie (Haavelmo 1944, 47).

To epistemiczne podejście do uwzględniania prawdopodobieństwa w modelach przyczynowości zilustruję hipotetycznym przykładem. Interesującą badawczo wielkością jest Y , czyli konsumpcja produktu A przez daną jednostkę. W hipotetycznym modelu jest proponowanych n czynników X_1, \dots, X_n wpływających na tę konsumpcję, zgodnie z funkcją $Y = f(X_1, \dots, X_n)$. Spośród tych czynników dla danej osoby uwzględnia się $i \leq n$. Jak zauważa Haavelmo, „jeśli postępujemy tak dla każdej jednostki, to przekonamy się, że niezależnie od tego, jaka jest funkcja f , zaproponowane «wyjaśnienie» będzie niekompletne” (1944, 50).

Nawet bowiem w tych samych okolicznościach (dla tych samych wartości zmiennych X_1, \dots, X_i) dana jednostka może mieć różne wartości konsumpcji Y tego samego produktu. Haavelmo podkreśla, że

Możemy spróbować usunąć te rozbieżności, wprowadzając więcej „czynników wyjaśniających” X . Zwykle jednak szybko wyczerpiemy listę takich czynników, które można by uznać za *wspólne* wszystkim jednostkom, a które w tym samym czasie miałyby niezaniebdywalny wpływ na Y . Rozbieżności [między faktyczną a przewidywaną na podstawie modelu konsumpcją] dla każdej jednostki mogą zależeć od wielkiej liczby czynników, które mogą być różne dla poszczególnych jednostek oraz mogą zmieniać się w czasie dla każdej jednostki. (1944, 50)

Zasadniczo więc, jak sugeruje Haavelmo, podanie w pełni *deterministycznego* modelu zależności przyczynowych między konsumpcją a „czynnikami ją wyjaśniającymi” byłoby możliwe dla poszczególnych jednostek i określonego okresu czasu: „równanie stochastyczne dla n zmiennych może być dokładnym [tj. deterministycznym] równaniem dla $n + m$ zmiennych” (1944, 56).

Równania stochastyczne natomiast można przekształcić do postaci równań deterministycznych, nadając zmiennym wyjaśniającym X nową interpretację. Pierwszy rodzaj takich równań deterministycznych opiera się na założeniu, że nie ma „błędów w równaniach” (Qin i Gilbert 2001, 430-32). Otrzymuje się je, przypisując błędom losowym w poszczególnych równaniach modelu strukturalnego stałą wartość, np. zera. Każda ze zmiennych wyjaśniających X jest interpretowana wówczas „jakby nie było błędu”.

Drugi rodzaj równań deterministycznych, wyprowadzanych ze „stochastycznych”, to równania dla wartości oczekiwanej. Reprezentują one średnie zależności między zespołami obserwacji. Nie upraszcza się zatem równań w modelu wyjściowym, lecz bada ich własności graniczne.

Taką możliwość, polegającą na podawaniu kompletnej listy czynników dla każdej jednostki w danym okresie czasu, odrzuca stanowczo nie tylko

Haavelmo, ale i Rubin. W modelach RCM ten drugi proponuje taką modyfikację listy, która pozwoli uwzględnić faktyczne zależności przyczynowe na poziomie jednostkowym bez wprowadzania prawdopodobieństwa epistemicznego. Zauważmy przy tym, że model potencjalnych rezultatów Rubina zmieniałby jednak pierwotne pytanie o to, jakie są przyczyny takiej, a nie innej zaobserwowanej konsumpcji danego produktu przez jednostkę, na pytanie o to, jakie są różnice w tej konsumpcji w sytuacjach, gdy na tę jednostkę oddziałują różne przyczyny.

Konsekwencją przyjętego przez Haavelma rozwiązania jest nieczytelność interpretacji zależności między zmiennymi w modelu, wynikająca bezpośrednio z:

konwencji, że uwzględnione [w modelu] zmienne obserwowalne są złożone z dwóch części, a mianowicie, *systematycznej* części, która z założenia spełnia dokładnie przyjętą zależność oraz z części błędowej (ang. error part) czy „zakłócenia”, które ma naturę stochastyczną”. (1944, 52-53)

Dodatkową komplikacją przy podaniu jednoznacznej, a zwłaszcza przyczynowej, interpretacji tak określonych w modelu strukturalnym zależności jest to, że „takie rozdzielanie zmiennych jest z konieczności względne, zależąc od określonego systemu równań teoretycznych, które rozpatrujemy” (1944, 53).

Uzasadnieniem stosowania takich modeli o niejednoznacznie określonej interpretacji jest – zdaniem Haavelma – to, że

mamy bogaty zestaw przypadków pokazujących, że takie schematy dają się w udany sposób stosować do obserwowanych zjawisk, a po drugie, ponieważ teoria takich schematów jest rozwinięta lepiej niż jakichkolwiek innych schematów aproksymacyjnych. (1944, 56)

Dodatkowym uzasadnieniem ich stosowania do zjawisk ekonomicznych jest fakt, że zwykle mamy do czynienia tylko z całkowitymi lub średnimi efektami wielu jednostkowych decyzji, które częściowo są motywowane takimi samymi czynnikami, ale częściowo czynnikami specyficznymi dla poszczególnych jednostek. (1944, 56)

Takie uzasadnienie dla współczesnych ekonomistów jest niezadowolające (Heckman 2000), chociaż z powodów innych niż wskazane wyżej ze stanowiska proceduralnego.⁵¹ Zasadniczą trudność podejścia do modelowa-

⁵¹ Do trwałych owoców pracy Komisji Cowlesa Heckman zalicza wyłącznie wprowadzenie pojęć: zmiennej egzogenicznej (zewnętrznej) i endogenicznej (wewnętrznej) oraz oma-

nia zjawisk ekonomicznych, jakie zaproponowała Komisja Cowlesa, Heckman dostrzega w tym, że posługiwała się ona globalnymi danymi dla rynku oraz stosowała keynesowskie podejście w tworzeniu systemów ogólnej równowagi.⁵²

Ilustracją, która szczegółowo wykazuje takie błędy w pracach inspirowanych teorią Komisji Cowlesa, jest podjęta przez T. Liu (1960) krytyka wpływowego modelu Kleina-Goldbergera (1955) ekonomii USA. Podważa on stosowane tam zasady identyfikacji, które w sposób sztuczny prowadzą do włączenia do modeli więcej zmiennych niż w innych ówczesnie spotykanych modelach. W konsekwencji powstaje problem, na który – jak zaznacza Liu – nie zwracano od czasu prac Haavelma uwagi, a mianowicie problem niewystarczającej ilości danych do identyfikacji parametrów, czyli problem niedoidentyfikowania.⁵³ W dyskutowanym modelu występuje dwadzieścia zmiennych endogenicznych oraz piętnaście egzogenicznych, konsumpcja jest funkcją sześciu, inwestycje – trzech. Liu proponuje (1960, 857), by zamiast uwzględniać, np. w przypadku inwestycji, tylko te zmienne, które spełniają określone kryteria statystyczne (takie, jak „rozsądne” czy „istotne” współczynniki regresji), brać pod uwagę te zmienne, które są uznane za istotne na podstawie teorii ekonomicznej. Wówczas w tym modelu Kleina-Goldbergera inwestycja byłaby funkcją siedemnastu zmiennych, co pokazuje jednocześnie niedoidentyfikowanie parametrów strukturalnych. Ogólna krytyka, jaką kieruje Liu pod adresem modelu Kleina-Goldbergera, dotyczy więc znacznie szerszej klasy modeli, w tym także współcześnie stosowanych modeli sieci bayesowskich:

Kryterium, jakie Klein i Goldberger stosują przy selekcji lub odrzuceniu zmiennych wyjaśniających, rzuca dużo światła na zasady, jakie doprowadziły ich do wyspecyfikowania takiej struktury. Właśnie te zasady są odpowiedzialne za nadidentyfikację strukturalnych zależności Kleina-Goldbergera. Zasadniczo istotna zmienna wyjaśniająca jest wykluczona [poza model], jeśli włączenie jej doprowadziłoby do istotnie nierozsądnej wielkości jej współ-

wianych tu jako „autonomiczne” parametrów strukturalnych, niezmienniczych względem polityki ekonomicznej (Heckman 2000, 46).

⁵² Jedną z późniejszych reakcji na problemy modelowania strukturalnego, jakie dostrzegano od lat 1960., było definiowanie – na niższym poziomie analizy – parametrów strukturalnych w kategoriach preferencji oraz parametrów technologicznych (Heckman 2000, 48).

⁵³ Ten problem kontrastuje z powszechnie dyskutowanym, jak zaznacza Liu, a występującym jeszcze w pracy Wrighta (1921), problemem nadidentyfikacji parametrów – większej ilości równań identyfikujących parametry modelu niż same parametry. Wright wykorzystywał nadidentyfikację do testowania modeli (Goldberger 1972).

czynnika lub innych współczynników strukturalnych. Skrętnie prowadzone eksperymenty z alternatywnymi kombinacjami z coraz to mniejszą i mniejszą liczbą zmiennych wyjaśniających doprowadziły ostatecznie do zbioru „istotnych” i „rozsądnych” współczynników. Tak otrzymany rezultat jest jednak *tylko estymacją struktury*. (Liu 1960, 858; podkreślenie – P.K.)

W miejsce takiej, zdaniem Liu, niedopuszczalnej teoretycznie i obliczeniowo, procedury omijania ograniczeń w dostępnych danych w celu podania zależności strukturalnych (przyczynowych) proponuje on tworzenie modeli uwzględniających wszystkie zależności umotywowane teoretycznie, a następnie wyprowadzenie drogą jednorównaniowych regresji najmniejszych kwadratów zmiennych łącznie zależnych od zmiennych egzogenicznych i wykorzystanie tego równania w predykcjach (1960, 860).

Jak podkreśla Heckman, w tym samym czasie uznano modele tworzone zgodnie z podejściem Komisji Cowlesa za „empiryczną porażkę”:

Niestabilność estymowanych parametrów oraz praktyka ponownego dopasowania makromodeli w krótkich okresach czasu wskazywała, że oficjalna retoryka ekonometrii nie była stosowana w praktyce. (2000, 71)

Heckman opisuje dwie zasadniczo różne reakcje, związane z niepowodzeniem programu Komisji Cowlesa (2000, 72). Pierwsza z nich jest kontynuacją tego podejścia do modelowania, chociaż posługuje się już „jawniej sparametryzowanymi” modelami. Druga reakcja, zainicjowana przez artykuł C. Grangera (1969) oraz prace C. Simsa (1972; 1977), to wprowadzenie modeli autoregresji wektorowej (VAR).⁵⁴

Heckman podstawową różnicę między tymi podejściami dostrzega w odniesieniu do danych. Podejście strukturalistyczne nadaje priorytet teorii, nawet za cenę niedobrego dopasowania do danych. Podejmuje się tu jednak pytania o kontrfaktyczne efekty polityki ekonomicznej i dokonuje analiz zależności przyczynowych. Podejście VAR natomiast jako priorytetowe traktuje dane zsumowane w postaci wektora autoregresji, a teorię ekonomiczną tylko jako nieformalną wskazówkę przy dokonywaniu statystycznych dekompozycji danych.

W odniesieniu do podejścia VAR Heckman (2000, 72-74) krytycznie zaznacza, że podsumowywanie danych w postaci wektorów autoregresji jest nieprzejrzyste dla niewtajemniczonych w stosowane przy tym podejściu

⁵⁴ Por. (Box i Jenkins 1976; Enders 1989; Hendry i Mizon 2000; Jadamus-Hacura 2002; Granger i Jeon 2003).

techniki. Ponadto nie unika się tu problemów z identyfikowalnością, które wcześniej krytykowano w odniesieniu do programu Komisji Cowlesa.⁵⁵

odwołanie się do współczesnych metod [analizy] szeregów czasowych wydaje się zastępować czarną skrzynkę identyfikacji Cowlesa czarną skrzynką identyfikacji w oparciu o szeregi czasowe. (2000, 74)

Zmodyfikowane podejście strukturalne zmierza do uniknięcia problemu czarnej skrzynki przy identyfikacji modeli poprzez odwołanie się do danych na poziomie mikro (zwłaszcza dane panelowe i przekroje) przy wprowadzaniu interpretacji i ograniczeń parametrów strukturalnych w modelach.⁵⁶ Jako wzorcowy dla wczesnych prac Heckman przytacza model popytu na kapitał oraz model podaży na pracę kobiet. Oba modele interpretują, jego zdaniem, zjawiska ekonomiczne w sposób klarowny, z uwzględnieniem zasadniczych idei teoretycznych w ekonomii.

Interpretowalność parametrów oraz możliwość wyprowadzania kontrfaktycznych przewidywań (uwzględniających zmiany dokonane w polityce ekonomicznej) nasiliły stosowanie modeli strukturalnych w latach 1970. i 1980. zarówno w badaniach makro-, jak i mikroekonomicznych. Zasadniczy wkład tych prac, zdaniem Heckmana, polega na tym, że pozwalają one ocenić, jakie informacje można uzyskać z danych, pod warunkiem uprzedniego przyjęcia określonych założeń apriorycznych. A postęp w stosunku do podejścia Komisji Cowlesa polega przede wszystkim na interpretacji parametrów strukturalnych poprzez odniesienie do „preferencji, technologii oraz oczekiwań”, a także na dyskusji identyfikowalności tych parametrów w zróżnicowanych kontekstach empirycznych (2000, 76).

Postęp, jaki dokonał się od czasu zakończenia prac Komisji Cowlesa, nie jest jednak zbyt wielki:

empiryczne sukcesy strukturalnego programu badawczego nie są imponujące. [...] Na ogół estymowalne są tylko najprostsze modele strukturalne, a te często odrzucane są przez dane, co szczególnie podkreślają zwolennicy programu VAR. Estymowane są nieliczne modele strukturalne, pozwalające systematycznie wyjaśnić szeroką gamę działań ekonomicznych, na które napotykają analizowane podmioty. (Heckman 2000, 76)

⁵⁵ Tę krytykę rozwija (Hansen 2005).

⁵⁶ Heckman podkreśla rolę w udostępnianiu danych z poziomu mikro przez Instytut Badań Społecznych Uniwersytetu w Michigan (Heckman 2000, 74).

Ponadto, w praktyce w celu umożliwienia estymacji modeli mających przewidywać skutki polityki ekonomicznej, często ignoruje się tę politykę oraz wprowadza zależności funkcyjne niezgodne z danymi.⁵⁷

Przykładu typowej negatywnej reakcji ekonomistów na podejście strukturalne do modelowania zależności między zjawiskami ekonomicznymi dostarcza praca G. Lewisa (1986). Badając wpływ przyczynowy przynależności do związku zawodowego na wysokość płacy, Lewis pokazał, że modele strukturalne dopuszczają liczne estymacje parametrów przy zastosowaniu różnych metod oraz różnych zmiennych warunkujących. Ta zmienność była powodem odrzucenia przez niego modeli strukturalnych i powrotu do metod najmniejszych kwadratów, które dają lepsze (bardziej odporne) estymacje.

Komentując pracę Lewisa i zgadzając się z jego krytyką nieprzejrzystości parametrów (czarne skrzynki), Heckman podkreśla, że

wykazana przez Lewisa wrażliwość estymacji parametrów przyczynowych na alternatywne założenia identyfikujące nie dowodzi niemożliwości estymacji efektu przyczynowego związków zawodowych na płace. (2000, 78)

Różne metody estymacji, zdaniem Heckmana, doprowadzą do takich samych rezultatów estymacji parametrów przyczynowych w modelu strukturalnym tylko pod warunkiem, że nie mamy do czynienia z obciążeniem selekcją. Sugerowanym przez niego wyjściem jest m.in. zastosowanie statystycznych metod estymacji nieparametrycznej oraz analizy wrażliwości. Zasadnicze jednak rozwiązanie w unikaniu wieloznaczności wyników w modelowaniu strukturalnym ma polegać, zdaniem Heckmana rozwijaniu teorii ekonomicznej (2000, 79-82).

4.4 Modele sieci bayesowskich (BN)

W ciągu ostatnich dwóch dekad opracowano podstawy reprezentacji zależności przyczynowych, o których twierdzi się, że są uogólnieniem dotychczas stosowanych metod modelowania zależności przyczynowych. Dotyczy to omówionych wcześniej modeli RCM, a przede wszystkim modeli równań

⁵⁷ Stosowanie rozwiniętych w latach 1980. i 1990. technik estymacji Eulera, pozwalających na uproszczoną estymację wybranych parametrów modelu, Heckman traktuje jako wyraz całkowitego zarzucenia pierwotnego programu Komisji Cowlesa, zmierzającego do estymacji efektów polityki makroekonomicznej (Heckman 2000, 77). Por. też (Hendry i Clements 2003).

strukturalnych SEM. Z tego też względu temu właśnie podejściu do modelowania zależności przyczynowych poświęcę w tej pracy więcej miejsca (por. też Dodatki D-G).

Zasadnicza nowość tego podejścia polega na wykorzystaniu opracowanych w końcu lat 1980. tzw. sieci bayesowskich (Pearl 1986; 1988) do reprezentacji graficznej zależności przyczynowych. Sieci bayesowskie, na które składają się węzły oraz ich połączenia (najczęściej zorientowane jednokierunkowo), dostarczają nowej notacji matematycznej dla reprezentowania asymetryczności relacji przyczynowych, a jednocześnie zawierają rekurencyjny mechanizm uaktualniania zależności w oparciu o nowe dane. Jak zaznacza S. Lauritzen, jeden z nielicznych statystyków zaangażowanych od początku w opracowywanie podstaw reprezentacji przyczynowej za pomocą sieci bayesowskich,

Stало się jasne, że modele graficzne, a w szczególności oparte na skierowanych grafach acyklicznych, mają naturalną interpretację przyczynową i w ten sposób tworzą podstawę języka, w którym można precyzyjnie analizować i dyskutować pojęcia przyczynowe. (2001, 63)

Po wstępnym wprowadzeniu podstawowych zagadnień związanych z modelowaniem zależności przyczynowych za pomocą sieci bayesowskich, przejdę do precyzyjniejszej prezentacji podstaw tego podejścia, a następnie odniosę je do innych sposobów modelowania.

Prezentację podejścia BN do modelowania zależności przyczynowych rozpocznę od krótkiego przedstawienia pojęć dotyczących grafów, a w szczególności zorientowanych grafów acyklicznych.⁵⁸ *Grafem* jest para $G = (V, E)$, gdzie V to skończony zbiór węzłów w grafie, a E – podzbiór zbioru uporządkowanych par węzłów $V \times V$, czyli zbiór krawędzi łączących pary różnych zmiennych (wyklucza się pętle).⁵⁹ Węzły V zwykle interpretuje się jako zbiór zmiennych, a krawędzie (albo ich brak) jako odpowiednik zależności przyczynowej między tymi zmiennymi (albo jej brak).

Zbiór wszystkich zmiennych, które są połączone krawędziami z daną zmienną, np. α , jest określany jako zbiór zmiennych z nią *sąsiadujących*: $ne(\alpha)$.

⁵⁸ Por. też (Whittaker 1990; Jensen 1996; Lauritzen 1996; Cichosz 2000; Dawid 2001; Wierzchoń i Kłopotek 2002).

⁵⁹ W przypadku, gdy oprócz *reprezentacji* zależności przyczynowych realizowane jest zadanie *odkrywania* takich zależności na podstawie danych, podane definicje należy uzupełnić o dodatkowe rodzaje połączeń między zmiennymi. Por. też przypis kolejny.

W przypadku, gdy dla węzłów α oraz β tylko jedna uporządkowana para, np. $\langle \alpha, \beta \rangle$, jest elementem E , a więc gdy $\langle \alpha, \beta \rangle \in E$ oraz $\langle \alpha, \beta \rangle \notin E$, wówczas taka krawędź jest skierowana jednostronnie, co będę zapisywał też jako $\alpha \rightarrow \beta$ i interpretował jako asymetryczną zależność przyczynową między tymi węzłami. W takim przypadku α jest określane jako *rodzic* β , a β jako *dziecko* α . Zbiór rodziców danego węzła α oznacza się jako $pa(\alpha)$, a zbiór jego dzieci jako $ch(\alpha)$.

Jeśli do zbioru krawędzi nie należy żadna z par: $\langle \alpha, \beta \rangle \notin E$ oraz $\langle \alpha, \beta \rangle \in E$, wówczas mówi się, że dane węzły α i β nie są *połączone*.

Te określenia uogólnia się w naturalny sposób na zbiory węzłów $A \subset V$. I tak, rodzice A to zbiór wszystkich rodziców węzłów w A z wykluczeniem samych węzłów A : $pa(A) = \cup_{\alpha \in A} pa(\alpha) \setminus A$. Podobnie dla dzieci oraz dla zbioru zmiennych sąsiadujących z A : $ch(A) = \cup_{\alpha \in A} ch(\alpha) \setminus A$ oraz $ne(A) = \cup_{\alpha \in A} ne(\alpha) \setminus A$.

Jeśli wszystkie krawędzie w grafie są skierowane,⁶⁰ wówczas mówi się o *grafie skierowanym*. Sieci bayesowskie, które mają reprezentować zależności przyczynowe, nie powinny zawierać cykli, czyli pętli, zaczynających się i kończących w tym samym węźle, ani skierowanych ścieżek, które prowadzą od danej zmiennej, przez inne i z powrotem do niej. Innymi słowy, podstawowy graficzny model zależności przyczynowych opiera się na acyklicznym grafie skierowanym (DAG – skrót ang. Directed Acyclic Graph). Jeśli wszystkie krawędzie w grafie są nieskierowane (do E należą wszystkie możliwe krawędzie), wówczas mówi się o grafie *nieskierowanym*. Graf nieskierowany, który powstał z G przez zastąpienie skierowanych krawędzi nieskierowanymi, oznaczamy przez G^- .

Granica danego węzła $bd(\alpha)$ jest zbiór jego rodziców oraz sąsiadów, natomiast granicą zbioru zmiennych $A \subset V$ jest taki podzbiór węzłów w $V \setminus A$, które są rodzicami lub sąsiadami węzłów w A : $bd(A) = pa(A) \cup ch(A)$. *Domknięciem* A jest $cl(A) = A \cup bd(A)$.

⁶⁰ Nie zawsze można podać kompletną specyfikację sieci bayesowskiej na podstawie dostępnych danych oraz wiedzy ogólnej. W dyskusji bardziej złożonych przypadków wprowadza się dodatkową reprezentację zależności w grafach, np. zmienne niemierzalne (ukryte) reprezentuje się jako pusty okrąg (co odróżnia je od zmiennych mierzalnych – pełnych kropek), a konfunder (nieuwzględniona w modelu zmienna, będąca wspólną przyczyną dwóch zmiennych w modelu) – jako dwukierunkowo zorientowaną krawędź oznaczoną linią przerywaną. W sytuacji, gdy ani dane doświadczenia, ani struktura grafu nie pozwala na jednoznaczny wniosek dotyczący zależności przyczynowej między dwiema zmiennymi, pozostawia się daną krawędź niezorientowaną (Spirtes i in. 2000, 140-52).

Podgrafem grafu $G = (V, E)$ nazywamy taki graf $G_A = (A, E_A)$, w którym $A \subseteq V$ oraz $E_A \subseteq E \cap (A \times A)$, a więc taki, który może zawierać tę samą liczbę węzłów, ale mniejszą liczbę krawędzi. Podgrafem G_A grafu G indukowanym przez zbiór węzłów A nazywamy taki podgraf G , w którym $E_A = E \cap (A \times A)$.

Ważnym pojęciem jest także pojęcie grafu *kompletnego*, a więc takiego, w którym wszystkie węzły są połączone krawędziami. Analogicznie odnosi się to do podzbioru węzłów danego grafu – jest on kompletny, jeśli indukuje kompletny podgraf danego grafu. Taki podgraf nazywany jest *kliką*, jeśli jest maksymalny, a więc jeśli nie jest podzbiorem właściwym innego kompletnego podgrafu.

Ścieżką (o długości n) z węzła α do węzła β w grafie G jest sekwencja różnych węzłów $\alpha_0, \dots, \alpha_n$ taka, że dla dowolnego $i = 1, \dots, n$ $\alpha_i > \alpha_{i-1}$, $\alpha_i > \alpha_{i-1}$. W przypadku, gdy dopuści się możliwość, że w ścieżce $\alpha_0 = \alpha_n$, wówczas otrzymujemy cykl. Graf skierowany, który nie zawiera cykli, określany jest jako *acykliczny graf skierowany*.

Ścieżka w nieskierowanej wersji G^- grafu G jest określana jako *szlak*.

W skierowanym grafie acyklicznym zbiór wszystkich węzłów, dla których zachodzi $\alpha \rightarrow \beta$, ale nie $\beta \rightarrow \alpha$, określany jest jako zbiór przodków β , czyli $\text{an}(\beta)$. Natomiast zbiór wszystkich węzłów, dla których zachodzi $\alpha \rightarrow \beta$, ale nie $\beta \rightarrow \alpha$, określany jest jako zbiór potomków α $\text{an}(\alpha)$. Zbiorem węzłów $\text{nd}(\alpha)$, które nie są potomkami α , jest zbiór $V \setminus (\text{de}(\alpha) \cup \alpha)$. Jeśli podzbiorem zbioru A jest zbiór rodziców węzłów w A , to A określa się jako zbiór poprzedników A .

Podzbiór zmiennych $C \subseteq V$ jest *separatorem* dla węzłów α i β , jeśli wszystkie szlaki z α do β przecinają C . Jeśli dla wszystkich węzłów w A i B podzbiór C jest separatorem, to mówi się, że *separuje* A od B .

Moralny nieskierowany graf G^m otrzymuje się ze skierowanego acyklicznego grafu G przez: 1) połączenie niezorientowanymi krawędziami wszystkich par węzłów mających wspólne dzieci, a następnie 2) podanie nieskierowanej wersji otrzymanego grafu.

Zasadniczą kwestią przy reprezentacji rozkładu prawdopodobieństwa w grafie jest jego odniesienie do prawdopodobieństwa warunkowego. Każdemu z węzłów odpowiada zmienna losowa $(X_\alpha)_{\alpha \in V} = X_V$ o wartościach $(x_\alpha)_{\alpha \in V}$.

W przypadku zmiennych losowych X, Y oraz Z o rozkładzie łącznym P mówi się, że X jest *warunkowo niezależne* od Y warunkowo względem Z w rozkładzie $P - X \perp\!\!\!\perp Y \mid Z [P]$ – jeśli dla każdego mierzalnego zbioru w prze-

strzeni próbek dla X istnieje wersja prawdopodobieństwa warunkowego $P(A|Y, Z)$, która jest funkcją samego Z . Dla zmiennych dyskretnych warunek ten przyjmuje postać: $P(X = x, Y = y|Z = z) = P(X = x|Z = z)P(Y = y|Z = z)$.

Gdy te trzy zmienne dopuszczają łączną gęstość w odniesieniu do miary produktowej μ , otrzymujemy:

$$X \perp\!\!\!\perp Y | Z \text{ wtw } f_{X\eta Z}(x, y|z) = f_{X|Z}(x|z) f_{\eta Z}(y|z). \quad (4.34)$$

Dla dowolnej funkcji mierzalnej h dla przestrzeni próbek X własności relacji trójargumentowej „ $X \perp\!\!\!\perp Y | Z$ ” są następujące:

$$\text{jeśli } X \perp\!\!\!\perp Y | Z, \text{ to } Y \perp\!\!\!\perp X | Z, \quad (C1)$$

$$\text{jeśli } X \perp\!\!\!\perp Y | Z \text{ oraz } U = h(X), \text{ to } U \perp\!\!\!\perp Y | Z, \quad (C2)$$

$$\text{jeśli } X \perp\!\!\!\perp Y | Z \text{ oraz } U = h(X), \text{ to } X \perp\!\!\!\perp Y | (Z, U), \quad (C3)$$

$$\text{jeśli } X \perp\!\!\!\perp Y | Z \text{ oraz } X \perp\!\!\!\perp W | (Y, Z), \text{ to } X \perp\!\!\!\perp (W, Y) | Z, \quad (C4)$$

$$\text{jeśli } X \perp\!\!\!\perp Y | Z \text{ oraz } X \perp\!\!\!\perp Z | Y, \text{ to } X \perp\!\!\!\perp (Y, Z). \quad (C5)$$

Własność (C5) nie jest uniwersalna, lecz zachodzi tylko przy spełnieniu warunku nietrywialnej logicznej zależności między Y a Z .

Dla funkcji gęstości prawdopodobieństwa f z własności relacji niezależności probabilistycznej (C1)–(C4) wynikają następujące równości:

$$X \perp\!\!\!\perp Y | Z \text{ wtw } f(x, y, z) = f(x, z) f(y, z) / f(z) \quad (4.35)$$

$$X \perp\!\!\!\perp Y | Z \text{ wtw } f(x|y, z) = f(x|z)$$

$$X \perp\!\!\!\perp Y | Z \text{ wtw } f(x, z|y) = f(x|z) f(z|y)$$

$$X \perp\!\!\!\perp Y | Z \text{ wtw } f(x, y, z) = h(x, z) k(y, z) \text{ dla pewnych } h \text{ oraz } k$$

$$X \perp\!\!\!\perp Y | Z \text{ wtw } f(x, z, y) = f(x|z) f(y, z).$$

Używając A na oznaczenie X_A , otrzymujemy następujące własności relacji niezależności warunkowej dla zbiorów zmiennych:

$$\text{jeśli } A \perp\!\!\!\perp B \mid C, \text{ to } B \perp\!\!\!\perp A \mid C, \quad (C1')$$

$$\text{jeśli } A \perp\!\!\!\perp B \mid C \text{ oraz } D \subseteq B, \text{ to } A \perp\!\!\!\perp D \mid C, \quad (C2')$$

$$\text{jeśli } A \perp\!\!\!\perp B \mid C \text{ oraz } D \subseteq B, \text{ to } A \perp\!\!\!\perp B \mid C \cup D, \quad (C3')$$

$$\text{jeśli } A \perp\!\!\!\perp B \mid C \text{ oraz } A \perp\!\!\!\perp D \mid B \cup C, \text{ to } A \perp\!\!\!\perp B \cup D \mid C, \quad (C4')$$

$$\text{jeśli } A \perp\!\!\!\perp B \mid C \cup D \text{ oraz } A \perp\!\!\!\perp C \mid B \cup D, \quad (C5')$$

to $A \perp\!\!\!\perp B \cup C \mid D$ dla rozłącznych zbiorów A, B, C i D .

Lauritzen (2001, 68) proponuje następujące intuicyjne odczytanie zależności „ $X \perp\!\!\!\perp Y \mid Z$ ” w kontekście lektury książek: gdy znana jest książka Z , lektura książki Y jest zbędna przy czytaniu książki X .⁶¹

Własności niezależności warunkowej dla rozkładów łącznych zbiorów zmiennych losowych opisuje się zwięźle jako własności Markowa. Poniżej podaję kilka podstawowych:

- (P) własność Markowa dla par jest spełniona dla dowolnej pary niesąsiadujących węzłów α i β , jeśli $\alpha \perp\!\!\!\perp \beta \mid V \setminus \{\alpha, \beta\}$,
- (L) lokalna własność Markowa jest spełniona, jeśli dla dowolnego węzła α zachodzi, że $\alpha \perp\!\!\!\perp V \setminus \text{cl}(\alpha) \mid \text{bd}(\alpha)$,
- (G) globalna własność Markowa jest spełniona, jeśli dla dowolnych trzech (A, B, S) rozłącznych podzbiorów V takich, że S separuje A od B , zachodzi $A \perp\!\!\!\perp B \mid S$.

⁶¹ Warunek (C2) czyta się więc: „Gdy znana jest książka Z , wówczas lektura książki Y jest zbędna przy lekturze książki X , a więc lektura Y jest zbędna przy lekturze dowolnego rozdziału U książki X ”. (C3): „Gdy znana jest Z , wówczas lektura Y jest zbędna przy lekturze X , a więc lektura Y będzie też zbędna po lekturze dowolnego rozdziału U książki X ”. (C4): „Gdy znana jest Z , wówczas lektura Y jest zbędna przy lekturze X , i gdy nawet po przeczytaniu Y zbędna przy lekturze X jest lektura W , wówczas lektura obu książek Y oraz W jest zbędna przy lekturze X ”.

Globalna własność Markowa jest istotna, gdyż określa ogólne kryterium dla decydowania, kiedy dwa zbiory zmiennych są niezależne warunkowo ze względu na trzeci zbiór zmiennych. Ponadto, nie można jej dalej wzmocnić.

Własności Markowa, podobnie jak niezależność warunkowa, są ściśle związane z faktoryzacją. W przypadku grafu G mówi się, że miara prawdopodobieństwa P ma faktoryzację wyznaczoną przez G , jeśli dla wszystkich zupełnych podzbiorów $\alpha \subseteq V$ (Lauritzen 2001, 66) istnieją nieujemne funkcje ψ_α , które zależą od x tylko poprzez x_α oraz istnieje miara produktowa $\mu = \otimes_{\alpha \in V} \mu_\alpha$ taka, że P ma gęstość f względem μ , gdzie f ma postać: $f(x) = \prod_{\alpha \text{ zupełny}} \psi_\alpha(x)$. Można przyjąć (Lauritzen 2001, 66), że tylko kliki są wybierane jako zupełne podzbiory $\alpha \subseteq V$: $f(x) = \prod_{c \in C} \psi_c(x)$, gdzie C oznacza zbiór klik grafu G . Jeśli P ma faktoryzację, mówimy, że ma własność Markowa (F).

Między wyróżnionymi tu własnościami Markowa zachodzą następujące zależności logiczne: (F) \rightarrow (G) \rightarrow (L) \rightarrow (P). W przypadku jednak, gdy P ma dodatnią gęstość, (P) implikuje (F), a więc wszystkie własności Markowa są równoważne.⁶² Także w przypadku, gdy spełniony jest warunek (C5'), równoważne są następujące własności Markowa: (G) \equiv (L) \equiv (P).⁶³

W przypadku grafów skierowanych D mówi się o rekursyjnej faktoryzacji (DF): $f(x) = \prod_{\alpha \in V} f(x_\alpha | x_{\text{pa}(\alpha)})$. Lauritzen (2001, 72) dowodzi twierdzenia, zgodnie z którym jeśli P ma własność (DF) wyznaczoną przez graf D , a więc jeśli ma wyznaczoną przez niego rekursywną faktoryzację, to ma faktoryzację wyznaczoną przez (nieskierowany) moralny graf D^m .

Z lokalnej własności Markowa dla grafu D^m wynika z kolei inna ważna własność, jaką jest niezależność węzła α od pozostałych węzłów w V warunkowo względem $\text{bl}(\alpha)$ (pokrycia Markowa dla α): $\alpha \perp\!\!\!\perp V \setminus \alpha \mid \text{bl}(\alpha)$. Pokrycie Markowa to zbiór węzłów sąsiadujących z α w moralnym grafie D^m . Pokrycie Markowa w wyjściowym grafie D to zbiór rodziców, dzieci oraz rodziców dzieci α : $\text{bl}(\alpha) = \text{pa}(\alpha) \cup \text{ch}(\alpha) \cup \{\beta: \text{ch}(\beta) \cap \text{ch}(\alpha) \neq \emptyset\}$.

Lauritzen (2001, 73) dowodzi, że jeśli P ma rekursywną faktoryzację wyznaczoną przez skierowany graf acykliczny D , wówczas $A \perp\!\!\!\perp B \mid S$ zawsze, ilekroć A i B są separowane przez S w $(D_{\text{An}(A \cup B \cup S)})^m$, czyli w moralnym grafie najmniejszego zbioru poprzedników zawierającego $A \cup B \cup S$. Te

⁶² Dowód podaje Lauritzen (1996).

⁶³ Por. (Pearl 1988).

własność określa się jako skierowaną globalną własność Markowa (DG). Podobnie jak dla globalnej własności Markowa, (DG) nie może być już wzmocnione i daje najmocniejszą regułę odczytywania relacji warunkowej niezależności z danego grafu skierowanego.

Verma i Pearl (1990) określają własność (DG), odwołując się do pojęcia szlaku między dwiema zmiennymi w D^- , czyli nieskierowanej wersji grafu D . Szlak π z a do b jest *zablokowany* przez S , jeśli zawiera węzeł $\gamma \in \pi$ taki, że (i) $\gamma \in S$ oraz strzałki na szlaku π nie stykają się główkami w γ , albo (ii) γ i wszyscy jego potomkowie nie są w S , a strzałki π stykają się główkami w węźle γ .⁶⁴

Szlak nie zablokowany przez S jest *aktywny*, a dwa zbiory węzłów grafu A oraz B są *d-separowane* przez S , jeśli wszystkie szlaki z A do B są zablokowane przez S . Lauritzen (1996, 48-49) wykazuje równoważność obu podejść: Dla dowolnych rozłącznych podzbiorów A , B oraz S skierowanego acyklicznego grafu D S d-separuje A od B wtedy i tylko wtedy, gdy S separuje A od B w $(D_{An(A \cup B \cup S)})^m$.

Lokalną skierowaną własność Markowa (DL) P ma wtedy, gdy dowolna zmienna jest warunkowo niezależna od zmiennych nie będących jej potomkami względem rodziców: $\alpha \perp\!\!\!\perp \text{nd}(\alpha) \mid \text{pa}(\alpha)$.

Własności (DF), (DG) oraz (DL) są równoważne pod warunkiem, że istnieje gęstość f (Lauritzen 2001, 75-76). Zbiór rozkładów prawdopodobieństwa, które spełniają te własności dla danego skierowanego grafu acyklicznego D , oznacza się jako $M(D)$.

O skierowanym acyklicznym grafie D mówi się, że jest *przyczynowy*⁶⁵ dla rozkładu prawdopodobieństwa P względem podzbioru zmiennych $B \subseteq V$, jeśli P ma własność Markowa względem P .⁶⁶

$$p(x) = \prod_{\alpha \in V} p(x_\alpha \mid x_{\text{pa}(\alpha)}). \quad (4.36)$$

⁶⁴ Geiger, Verma i Pearl (1990) pokazują, że kryterium nie da się wzmocnić, a Meek (1995) udowodnił ten rezultat dla skończonej przestrzeni stanów.

⁶⁵ Radykalną krytykę interpretacji przyczynowej sieci bayesowskich na gruncie subiektywnej interpretacji pojęcia prawdopodobieństwa przeprowadził F. Lad (1999). Szereg ograniczeń takiej interpretacji wyczerpał N. Cartwright (2004, 807).

⁶⁶ Inne podejście do reprezentowania zależności przyczynowych w acyklicznych grafach skierowanych opracował G. Shafer (1996), który efekt przyczynowy określa dla *zdarzeń*, a nie jak w omawianym podejściu dla *zmiennych losowych*. Por. też (Lauritzen 2001).

Przy omawianiu przyczynowych modeli Markowa, jednym z zasadniczych zagadnień jest wyróżnienie dwóch sposobów zmieniania (uaktualniania) rozkładu prawdopodobieństwa w modelu w odpowiedzi na otrzymane (nowe) informacje. Taka modyfikacja może mieć postać warunkowania w oparciu o *zaobserwowane* informacje. Może także mieć postać warunkowania wynikającego z *działania*, jakie podjęto w stosunku do zależności reprezentowanych w modelu, które modyfikuje ten model.

W pierwszym przypadku, gdy dzięki obserwacji uzyskujemy informację o zmiennej X (np. $X = x$), wówczas modyfikacja rozkładu innych zmiennych, np. Y , polega na podaniu rozkładu warunkowego: $p(Y = y | X = x) = p(y | x)$. Ten rodzaj warunkowania określany jest jako *warunkowanie przez obserwację*.

Inny rodzaj modyfikacji występuje w przypadku, gdy podjęto manipulację – zmieniającą zależności przyczynowe reprezentowane w systemie – która sprawia, że zmiennej X przypisana zostaje określona wartość, np. $X = x$.⁶⁷ Za Lauritzenem wprowadza się specjalną notację dla odróżnienia *warunkowania przez działanie lub interwencję*: $p(Y = y || X \leftarrow x) = p(y || x)$.⁶⁸

W niektórych przypadkach $p(y || x) = p(y | x)$, na ogół jednak wartości nowego rozkładu dla Y $p(y || x)$ nie można obliczyć na podstawie samej miary prawdopodobieństwa P bez dodatkowych założeń. Manipulacja bowiem wiąże się z modyfikacją zależności przyczynowych, które były reprezentowane w danym modelu i często w przypadku modeli BN używa się metafory „łamanie strzałek”. Mianowicie, jeżeli dokonuje się manipulacji na zmiennej X , to ta interwencja „łamię” wszystkie strzałki, które były skierowane do X . Dlatego też obliczanie efektu danej manipulacji dokonuje się w grafie zmodyfikowanym, który uwzględnia odpowiednie zmiany.

Za Lauritzenem zilustruję to na najprostszym przykładzie.

⁶⁷ To jest szczególnie przypadek manipulacji dokonanej na zmiennej, gdyż w ogólnym przypadku może ona prowadzić nie do przypisania konkretnej wartości, lecz do zmiany rozkładu zmiennej X , a nawet do nowego warunkowego rozkładu dla tej zmiennej, np. gdy w grupie eksperymentalnej uzależnia się rodzaj leczenia od odporności pacjentów (Spirtes i in. 2000, 50-53).

⁶⁸ Ze względu na to, że – jak zostanie to szczegółowiej omówione poniżej – efekty manipulacji i obserwacji są różne, ważna jest kolejność wykonywania poszczególnych operacji. Przyjmuje się kolejność operacji od prawej strony do lewej, np. $p(x || y | z) = P(X = x | Y \leftarrow y, Z = z)$ oznacza najpierw zaobserwowanie, że $Z = z$, a potem nadanie Y wartości y przez interwencję.



Rysunek 4.7.

Jeśli grafy przedstawione na Rysunku 4.7 są przyczynowymi grafami Markowa, to otrzymujemy następujące zależności. Dla grafu po lewej stronie nie ma znaczenia, czy dokonano interwencji, czy obserwacji nowej wartości zmiennej X : $p(y \parallel x) = p(y \mid x)$. Jest tak dlatego, że nie ma strzałek skierowanych do tej zmiennej, które „usunęłyby” interwencja. Inaczej jest w przypadku grafu po prawej stronie, gdzie: $p(x \parallel y) = p(x)$. W tym przypadku interwencja „łamie” strzałkę $X \leftarrow Y$ i zmodyfikowany rozkład zmiennej X po dokonaniu interwencji nie zależy już od Y .

Pearl stosuje w swoich pracach inną niż Lauritzen notację. Opiera się ona na regułach tzw. rachunku „zrób” (ang. ‘do’ calculus), które stosuje się do modyfikacji wyjściowego grafu przez usuwanie odpowiednich krawędzi. Z kolei Spirtes, Glymour i Scheines proponują dodanie do wyjściowego grafu tzw. zmiennej działania (ang. policy variable), reprezentującej podejmowaną interwencję. Zaletą ich podejścia jest to, że: (i) pozwala na reprezentowanie takich interwencji, które prowadzą nie tylko do ustalenia konkretnej wartości danej zmiennej, ale także takich, które prowadzą do zmiany jej rozkładu prawdopodobieństwa, także warunkowego; (ii) obliczanie efektu interwencji w zmodyfikowanym grafie nie wymaga dodatkowych reguł, poza opierającymi się na d-separacji (Spirtes, Glymour i Scheines 2000, 43-47).

Warunkowanie przez obserwację, zgodnie z twierdzeniem Bayesa, wyraża się dla przyczynowego grafu D następująco:

$$\begin{aligned}
 p(x \mid x_A^*) &= \frac{p(x)}{p(x_A^*)} \Big|_{x_A = x_A^*} = \\
 &= \frac{p(x)}{\sum_{y: y_A = x_A^*} p(y)} \Big|_{x_A = x_A^*} .
 \end{aligned}
 \tag{4.37}$$

W acyklicznym skierowanym przyczynowym grafie D spełniona jest także formuła, którą za Lauritzenem (2001, 77-78)⁶⁹ określa się jako *formułę interwencji*:

$$\begin{aligned}
 p(x \| x_A^*) &= \prod_{\alpha \in V \setminus A} p(x_\alpha | x_{\text{pa}(\alpha)}) \Big|_{x_A = x_A^*} = \\
 &= \frac{p(x)}{\prod_{\alpha \in A} p(x_\alpha^* | x_{\text{pa}(\alpha)})} \Big|_{x_A = x_A^*}.
 \end{aligned}
 \tag{4.38}$$

W przypadku tej formuły założenie o przyczynowym charakterze zależności sprowadza się do tego, że zależności warunkowe pozostają bez zmian dla tych zmiennych, których nie dotyczy interwencja.

Różnica między formułą interwencji (4.38) a formułą warunkowania przez obserwację (4.37) dotyczy mianownika, gdzie iloczyn specyfikacji warunkowych zastępuje prawdopodobieństwo brzegowe $p(x_A^*)$. Z tych formuł wynika więc, że gdy interwencji dokonuje się na jednej zmiennej, która w grafie D nie ma rodziców, wówczas obie formuły są równoważne: $p(x \| x_\alpha^*) = p(x | x_\alpha^*)$.

Przechodzimy z kolei do dyskusji na temat problemu identyfikowalności efektów interwencji w przypadku, gdy zaobserwowany jest tylko fragment systemu przyczynowego, a więc gdy tylko niektóre spośród zmiennych V są mierzalne. Załóżmy, że zmienne V mają rozkład łączny w acyklicznym grafie skierowanym D taki, że jest on przyczynowy w sensie Markowa względem D . W zbiorze V wyróżnia się dwa podzbiory: t – zmienne reprezentujące oddziaływanie oraz R – odpowiedź. Efekt przyczynowy t w stosunku do R wyraża formuła: $p(x_R \| x_t^*)$. Pozostałe zmienne w V należą bądź do grupy zmiennych zaobserwowanych O , w tym przypadku – współzmiennych, lub niezaobserwowanych U .

Zasadnicze w tym przypadku pytanie dotyczy tego, czy znany z badań rozkład łączny zmiennych t , R oraz O pozwala na określenie efektu przyczynowego t względem R , a więc czy na tej podstawie jest on identyfikowalny.

⁶⁹ Podobne formuły występują w tekstach innych autorów: por. (Robins 1986; Pearl 1993; Spirtes i in. 2000).

Za Lauritzenem identyfikowalność w tym przypadku definiuje się następująco:

O identyfikuje efekt przyczynowy t względem R , jeśli dla dowolnej pary rozkładów P_1, P_2 , które są przyczynowe w sensie Markowa, względem D zachodzi: $p_1(x_t, x_R, x_O) \equiv p_2(x_t, x_R, x_O) \rightarrow p_1(x_R \parallel x_t) \equiv p_2(x_R \parallel x_t)$.

Przy założeniu, że gdy O identyfikuje efekt przyczynowy t na R , można postawić pytanie, w jaki sposób określić minimalny zbiór współzmiennych (jako podzbiór O), wystarczający do zidentyfikowania tego efektu.

Z omówionych wcześniej sposobów reprezentowania efektów interwencji w dalszej części będą posługiwał się modyfikacją grafu przez dodanie zmiennej reprezentującej efekt interwencji. A więc jeśli interwencji dokonuje się na zmiennej X_α reprezentowanej w grafie przez węzeł α , wówczas dodany jest węzeł α' reprezentujący dodatkową zmienną F_α . Rozkład warunkowy zmiennej X_α w grafie „po-interwencyjnym” określa formuła:

$$p'(x_\alpha | x_{\text{pa}(\alpha)}, f_\alpha) = \begin{cases} p(x_\alpha | x_{\text{pa}(\alpha)}) \text{ dla wartości } f_\alpha \text{ spoza} \\ \text{zakresu } x_\alpha \\ 1 \text{ gdy } f_\alpha = x_\alpha^* \text{ oraz } x_\alpha = x_\alpha^* \\ 0 \text{ gdy } f_\alpha = x_\alpha^* \text{ oraz } x_\alpha \neq x_\alpha^* \end{cases} \quad (4.39)$$

Ustalenie zmiennych, które wystarczają do zidentyfikowania efektu interwencji, polega na takim przekształceniu odpowiedniej formuły zawierającej wyrażenia reprezentujące interwencje, by zawierała ona wyłącznie wyrażenia określające prawdopodobieństwa warunkowe. Te ostatnie bowiem – w przypadku mierzalnych współzmiennych – traktuje się jako obserwowalne.

Założmy zatem, że interwencja dotyczy węzła t , a pytanie – efektu na zmienne R . Wiedza obserwowalna dotyczy zmiennych O_A , które mają wartość o_A . Efekty tej interwencji rozważa się w zmodyfikowanym grafie D' , w którym dodano zmienną t' , reprezentującą interwencję. Ze skierowanej własności Markowa dla rozkładu P oraz P' można wyprowadzić następujące reguły wnioskowania (Lauritzen 2001, 83):⁷⁰

⁷⁰ Pearl (1995a) wysnuł hipotezę, że ten zbiór reguł jest kompletny, a więc że omawiany efekt przyczynowy można otrzymać wtedy i tylko wtedy, gdy stosuje się te reguły.

- obserwacja wartości⁷¹ t : jeśli R i t są d-separowane w D' warunkowo względem A , to $p(x_r | x_A, x_t) = p(x_r | x_A)$,
- interwencja⁷² na t : jeśli R i t' są d-separowane w D' warunkowo względem A , to $p(x_r | x_A || x_t) = p(x_r | x_A)$,
- równoważność obserwacji i interwencji⁷³ na t : jeśli R i t' są d-separowane w D' warunkowo względem $\{A, t\}$, to $p(x_r | x_A || x_t) = p(x_r | x_A, x_t)$.

Kryterium „tylnych drzwi”⁷⁴ Pearla określa, kiedy zbiór współzmiennych O pozwala na zidentyfikowanie efektu przyczynowego interwencji t w stosunku do R . Załóżmy, że $O_0 \subseteq O$ i O_0 spełnia warunki: (1) współzmiennych w O_0 pozostają poza zasięgiem interwencji (O_0 jest d-separowane w D' od t')⁷⁵ oraz (2) oddziaływanie interwencji ustala się po uwzględnieniu współzmiennych (R jest d-separowane w D' od t' warunkowo względem $O_0 \cup \{t\}$). Wówczas O identyfikuje efekt interwencji na t w stosunku do R następująco:⁷⁶

$$p(x_r || x_t^*) = \sum_{x_{O_0}} p(x_r | x_{O_0}, x_t^*) p(x_{O_0}). \quad (4.40)$$

Przykład sytuacji, gdy kryterium tylnych drzwi nie jest spełnione, ilustruje Rysunek 4.8.⁷⁷

⁷¹ Z tego, że R i t są d-separowane w D' warunkowo względem A , wynika, że są także d-separowane w D , gdyż wszystkie szlaki z t do R w D są także szlakami w D' . Skoro więc A blokuje szlaki w D' , to także w D . Z globalnej własności Markowa dla D wynika z kolei, że jeśli R i t są d-separowane w D' warunkowo względem A , to R i t są także niezależne warunkowo względem A .

⁷² Ta reguła wynika z faktu, że warunkowanie ze względu na interwencję na t w D jest równoważne obserwacji warunkowania na t' w D' .

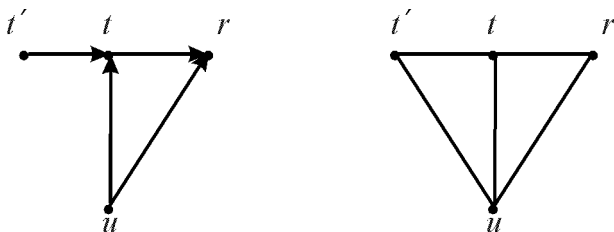
⁷³ Wynika podobnie jak w przypadku interwencji na t . Por. też twierdzenie 7.1 (Spirtes i in. 2000, 164).

⁷⁴ Nazwa „tylne drzwi” ma wskazywać na uwzględnienie wpływu współzmiennych na t , a więc takich szlaków od t do R , które rozpoczynają się w t w kierunku przeciwnym do skierowania (strzałki są skierowane w kierunku t , a nie R), a więc wychodzących poniekąd tylnymi drzwiami.

⁷⁵ Ten warunek można sformułować, żądając, by żadna zmienna w O_0 nie była potomkiem t . Wówczas można go sprawdzić w D , a nie w zmodyfikowanym grafie D' .

⁷⁶ To kryterium wynika z podanych tu reguł (Lauritzen 2001, 85).

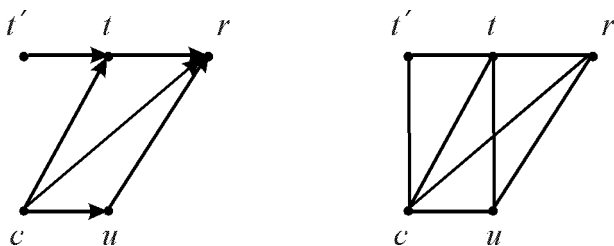
⁷⁷ Rozbudowany przykład znajduje się w Dodatku F.



Rysunek 4.8. (Opracowano na podstawie: Lauritzen 2001).

Oprócz interwencji, na t w tym przypadku oddziałuje jeszcze niezaobserwowany konfunder. Jak ilustruje to moralny graf, istnieje szlak od t' do r poprzez u , która nie przechodzi przez t , dlatego nie można zastosować kryterium tylnych drzwi. W konsekwencji nie można zidentyfikować efektu interwencji na t poprzez obserwację t oraz r .

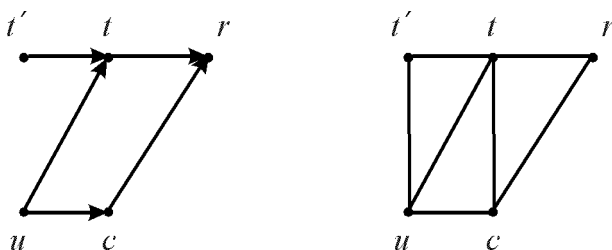
Problem konfundowania można rozwiązać, wprowadzając – w miejsce obserwacji t – randomizację oddziaływania, zależną od współzmiennnej c . Dzięki temu t jest niezależne od niezaobserwowanego konfundera u warunkowo względem c . Uwzględnienie zmiany związanej z randomizacją ilustruje Rysunek 4.9.



Rysunek 4.9. (Opracowano na podstawie: Lauritzen 2001).

Na podstawie grafu moralnego można w tym przypadku odczytać, że r i t' są w nim d-separowane warunkowo względem $\{c, t\}$, a tym samym spełnione jest kryterium tylnych drzwi. Efekt przyczynowy interwencji na t może więc być obliczony na podstawie danych dotyczących t , c oraz r .

Kolejne zastosowanie kryterium tylnych drzwi ma miejsce w badaniu nieeksperymentalnym, gdy znany jest zbiór współzmiennych wystarczający do ustalenia odpowiedzi. Jest to taki zbiór współzmiennych, że odpowiedź jest niezależna od niezaobserwowanej zmiennej warunkowo względem oddziaływania i współzmiennych.

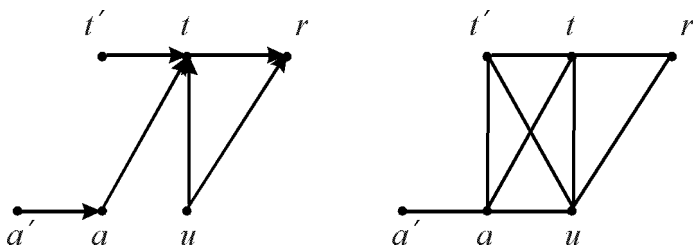


Rysunek 4.10. (Opracowano na podstawie: Lauritzen 2001).

W moralnym grafie, przedstawionym na Rysunku 4.10, można zauważyć, że spełnione są warunki zastosowania kryterium tylnych drzwi. Zmienna c oraz t' są d -separowane bezwarunkowo, a r i t' są d -separowane warunkowo względem zbioru $\{c, t\}$.

Inna często spotykana w badaniach nieeksperymentalnych sytuacja dotyczy *niepodatności*, a więc takiej sytuacji, gdy oddziaływanie jest przypisane pacjentom losowo, ale niektórzy z nich nie przestrzegają tego przypisania i są poddani oddziaływaniu innemu niż przypisane.

Niech t reprezentuje oddziaływanie faktycznie zrealizowane, natomiast a – przypisane. Graf dla tej sytuacji przedstawia Rysunek 4.11.



Rysunek 4.11. (Opracowano na podstawie: Lauritzen 2001).

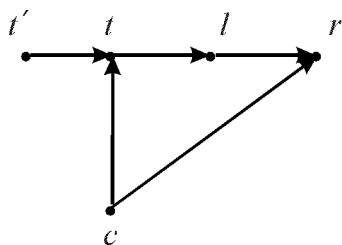
Zgodnie z grafem moralnym, r jest d -separowane od a' warunkowo względem $\{a, t\}$ i dzięki temu żądany efekt przyczynowy przypisanego oddziaływania jest identyfikowalny (ITT – skrót ang. intention-to-treat). Ze względu na to, że r nie jest d -separowane od t' , przez t nie można ocenić efektu faktycznie realizowanego oddziaływania.

Inny przypadek, kiedy współzmiennie pozwalają na zidentyfikowanie efektu przyczynowego, dotyczy takiej sytuacji, kiedy współzmiennie traktuje się jako aktywne czynniki determinujące odpowiedź. Tę sytuację określa podane przez Pearla kryterium drzwi frontowych.

Załóżmy, że $O_0 \subseteq O$ oraz że istnieje podzbiór $W \subseteq V \setminus (O_0 \cup R \cup \{t\})$ spełniający warunki: (1) zmienne w W pozostają poza zasięgiem interwencji (W jest d-separowane w D' od t'), (2) interwencja oddziałuje na współzmienną wyłącznie poprzez oddziaływanie niezależnie od zmiennych w W (O_0 jest d-separowane w D' od $(W \cup \{t'\})$ warunkowo względem t) oraz (3) interwencja oddziałuje na odpowiedź tylko poprzez współzmiennne po uwzględnieniu różnic z uwagi na zmienne w W (R jest d-separowane w D' od t' warunkowo względem $O_0 \cup W$). Wówczas O identyfikuje efekt interwencji na t w stosunku do R następująco:⁷⁸

$$p(x_R \| x_t^*) = \sum_{x_{O_0}} p(x_{O_0} | x_t^*) \sum_{x_t} p(x_R | x_{O_0}, x_t) p(x_t). \quad (4.41)$$

Rozważmy przypadek zilustrowany na Rysunku 4.12.



Rysunek 4.12. (Opracowano na podstawie: Lauritzen 2001).

W przypadku, gdy zaobserwowane są obie współzmiennne c oraz l , a także t i r , wówczas stosuje się kryterium tylnych drzwi dla c jako współzmiennnej i kryterium drzwi frontowych z l jako współzmienną.

Oprócz kwestii *przewidywania efektów interwencji* na zmienne w modelu przyczynowym, drugie podstawowe zadanie, jakie stawia się niekiedy bayesowskim modelom zależności przyczynowych, dotyczy *odkrywania zależności przyczynowych* na podstawie danych.⁷⁹ Wprowadza się tu drugie,

⁷⁸ Dowód podaje Lauritzen (2001, 89-90).

⁷⁹ Te prace zainicjował i kontynuuje zespół badawczy kierowany przez Glymoura w Instytucie Filozofii Uniwersytetu Carnegie Mellon w Pittsburghu. Szersze omówienie i odnośniki do literatury podaje Kawalec (2005).

obok własności Markowa, podstawowe dla modelowania przyczynowego BN założenie, jakim jest *założenie wierności* (ang. *faithfulness*).⁸⁰

Własność Markowa dotyczy tego, co wynika ze struktury grafu dla relacji warunkowej niezależności probabilistycznej. Drugie – jest poniekąd jego odwrotnością. Dotyczy mianowicie tego, co niezależności probabilistyczne zaobserwowane w danych implikują w stosunku do struktury grafu przyczynowego. Dlatego też założenie wierności jest podstawą wnioskowań, które na podstawie dostępnych założeń teoretycznych i wiedzy ogólnej oraz danych mają określić strukturę modelu przyczynowego w postaci skierowanego grafu acyklicznego lub zbioru grafów równoważnych obserwacyjnie, czyli wzorca.⁸¹

W wersji podstawowej założenie wierności określa się następująco (Spirtes i in. 2004, 457-59): jeśli graf D reprezentuje rozkład P dla zmiennych V , to P jest *wierne* względem D wtedy i tylko wtedy, gdy każda warunkowa niezależność, jaka zachodzi zgodnie z P , jest implikowana przez d-separację w D . Oznacza to, że taka niezależność warunkowa nie wynika tylko ze szczególnego układu wartości parametrów w rozkładzie, lecz że zachodzi dla wszystkich wartości wolnych parametrów rozkładu.⁸²

Intuicyjny sens założenia wierności określa, że mało prawdopodobne i niestabilne układy parametrów nie ukrywają rzeczywistego wpływu przyczynowego. W takim też sensie to założenie jest powszechnie przyjmowane w badaniach naukowych. Każdą bowiem teorię, która wyjaśnia określone zjawisko przez odwołanie się do bardzo szczególnego układu parametrów, traktuje się z podejrzeniem, jako sformułowaną *ad hoc*. Jeśli więc jakaś teoria postulowałaby, że trzy zmienne X oraz Z są niezwiązane przyczynowo przy wartości parametrów $\gamma = -(a\beta)$ i równaniach $Y = aX + \varepsilon_Y$, $Z = \beta X + \gamma Y + \varepsilon_Z$, wówczas taka teoria byłaby przedmiotem wątpliwości, gdyż

⁸⁰ Warunek ten można uznać za szczególną postać „założenia stabilności częstości”, którego rolę w badaniach empirycznych dyskutują M. Heller i J. Życiński (1987, 14-15).

⁸¹ Jak wspomniałem wcześniej, wzorec od grafu różni się przede wszystkim występowaniem krawędzi niezorientowanych. Grafy składające się na dany wzorec określa się po zorientowaniu każdej z takich krawędzi.

⁸² R. Giere dostarczył przykładu, który dobrze ilustruje pogwałcenie warunku wierności. Załóżmy, że w populacji ogólnej badamy zależność między intensywnymi ćwiczeniami fizycznymi a stanem zdrowia. Załóżmy, że duże obciążenie organizmu ćwiczeniami u połowy ćwiczących prowadzi do poprawy stanu zdrowia, a u drugiej połowy – do pogorszenia. Dla populacji ogólnej otrzymamy brak zależności między dwoma badanymi tu zjawiskami, ale wynika to wyłącznie ze szczególnej wartości parametrów, jakie wystąpiły w tym przypadku. Przykład podają za W. Strawińskim (2003, 231).

przy wszystkich innych wartościach tych parametrów spełnione byłoby założenie wierności, implikujące w tym przypadku zależności przyczynowe: $X \rightarrow Y$ oraz $X \rightarrow Z \leftarrow Y$.

Założenie wierności można podawać w różnych wersjach (Spirtes i in. 2004, 457-59). Jedna z nich dotyczy teorii wnioskowań bayesowskich: jeśli graf D nie implikuje, że dwie zmienne są nieskorelowane warunkowo względem trzeciej, to zbiór wartości wolnych parametrów, które implikują takie warunkowe nieskorelowanie ma prawdopodobieństwo pierwotne równe zero. Wersja założenia wierności, która nie wymaga przyjęcia istnienia prawdopodobieństwa pierwotnego, stwierdza, że: jeśli graf D nie implikuje, że dwie zmienne są nieskorelowane warunkowo względem trzeciej, to taka warunkowa korelacja jest niezerowa. Mocniejsza wersja założenia wierności podaje, że jeśli korelacja warunkowa jest słaba, to nie ma mocnej d-łączącej ścieżki⁸³ między tymi dwiema zmiennymi warunkowo względem trzeciej.

Dyskusję zagadnienia (zautomatyzowanego) odkrycia przyczyn⁸⁴ rozpoczną od najprostszego przypadku, gdzie zakłada się, że wszystkie zmienne ze zbioru V są mierzalne oraz że ten zbiór zmiennych jest przyczynowo *wystarczający*.⁸⁵ W przypadku, gdy graf przyczynowy jest znany, dla wielozmiennowych modeli o rozkładzie normalnym lub zmiennych dyskretnych istnieje jednorodnie zgodny MLE (maksymalny estymator wiarygodności, skrót ang. Maximum Likelihood Estimator)⁸⁶ współczynników krawędzi (Spirtes i in. 2004, 459-60). Otrzymuje się go przez regresję każdej zmiennej na rodzicach w przyczynowym skierowanym grafie acyklicznym, a dla zmiennych dyskretnych – przez wykorzystanie częstości względnej danej zmiennej warunkowanej względem rodziców tej zmiennej.

W przypadku, gdy nie przyjmuje się żadnego z założeń wierności, a graf nie jest znany, wówczas nie istnieje estymator kierunku strzałek dla zmiennych zależnych. Dowolny rozkład można bowiem reprezentować na podgrafie D , w którym każda para zmiennych sąsiaduje ze sobą niezależnie od kierunku strzałek.

⁸³ W sensie braku obecności na tej ścieżce d-separacji; por. (Spirtes i in. 2000, 14).

⁸⁴ Por. też (Gatnar 99-103).

⁸⁵ Zbiór V jest przyczynowo *wystarczający* dla danej populacji, jeśli każda przyczyna wspólna zmiennych z V jest w tym zbiorze lub ma taką samą wartość dla jednostek tej populacji (Spirtes i in. 2000, 22).

⁸⁶ Por. m.in. K. Janiszowski (2002, IV.39-IV.43).

Wprowadzenie jednak dodatkowych założeń, oprócz własności Markowa, sprawia, że dostępne są informacje wykluczające niektóre z możliwości, co z kolei pozwala – przynajmniej częściowo – przewidzieć efekty manipulacji. Zbiór wszystkich możliwych skierowanych acyklicznych grafów przyczynowych, które są zgodne z danymi, można opisać na różny sposób. Zasygnalizuję dwa z nich.

Grafy mogą być *równoważne ze względu na kryterium d-separacji*. Wówczas do klasy równoważnych obserwacyjnie modeli należą wszystkie te, które implikują te same zależności d-separacji, jakie stwierdza się w zaobserwowanych danych.

Można też charakteryzować równoważności z uwagi na rozkład prawdopodobieństwa. Dwa grafy D_1 oraz D_2 są *równoważne rozkładowo*, jeśli brzegowy rozkład prawdopodobieństwa dla zaobserwowanych zmiennych O , wygenerowany przez przypisanie wartości wolnym parametrom grafu D_1 , można również wygenerować przez przypisanie wartości wolnym parametrom grafu D_2 . Klasa równoważnych rozkładowo grafów daje opis dokładniejszy niż klasa grafów równoważnych ze względu na kryterium d-separacji, gdyż w ramach tej drugiej klasy mogą wystąpić modele, które nie są równoważne rozkładowo. Jednak z uwagi na złożoność obliczeń, niekiedy wybiera się ten pierwszy rodzaj równoważności, który wymaga posługiwania się wyłącznie graficznym kryterium d-separacji.

Jednym ze sposobów ustalenia równoważności rozkładowej jest odwołanie się do znanych już pojęć z teorii grafów, a ponadto do pojęcia *nieosłoniętego kolidera*: jest to taki węzeł X w grafie, że skierowane są do niego strzałki od dwóch innych niesąsiadujących ze sobą węzłów $A \rightarrow X \leftarrow B$.⁸⁷ Dla wielozmiennowych rozkładów normalnych oraz dla rozkładów dyskretnych dwa skierowane acykliczne grafy przyczynowe (bez skorelowanych błędów losowych) są *równoważne rozkładowo* wtedy i tylko wtedy, gdy mają te same węzły, te same zbiory zmiennych sąsiadujących i te same nieosłonięte kolidery (Verma, Pearl 1990; Spirtes i in. 2000, 305).

Podany powyżej sposób ustalania równoważności rozkładowej pozwala na reprezentowanie całej klasy modeli przyczynowych, które spełniają ten warunek. Otrzymujemy wówczas tzw. *wzorzec*. Wzorzec zawiera więc wszystkie węzły, nieosłonięte kolidery oraz krawędzie, które są nieskierowane, a skierowane pozostawia się tylko wówczas, jeśli występują w każdym modelu należącym do danej klasy.

⁸⁷ Wprowadzenie kolidera (por. Punkt 4.3) można przypisać S. Wrightowi.

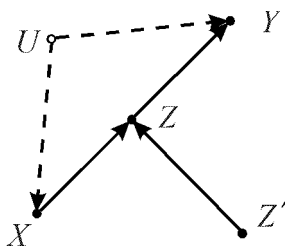
Algorytm PC programu TETRAD służy do odkrywania wzorców na podstawie danych. Jest on punktowo zgodnym estymatorem dla tych wzorców. Algorytm pobiera informacje o macierzy kowariancji lub danych dyskretnych, założenia dotyczące rozkładu oraz ewentualną wiedzę ogólną (dotyczącą np. uprzedniości czasowej). Poszukiwanie polega na obliczaniu niezależności warunkowych między zmiennymi, a wskazanie rodziców pozwala znacznie ograniczyć przestrzeń przeszukiwanych możliwości (Pearl 2000, 50; Spirtes, Glymour, Scheines 2000, 84-88).

Przeszukiwanie więzowe⁸⁸ ma tę zaletę, że nie wymaga estymacji wolnych parametrów modelu. Jednak jego wadą jest to, że daje w rezultacie jeden wzorec, nie sygnalizując innych wzorców, które wyjaśniałyby dane wprowadzone do algorytmu w podobnie zadowalający sposób, ale które zawierałyby inne modele przyczynowe.

Dodatkową komplikacją w poszukiwaniu modelu przyczynowego lub wzorca jest wprowadzenie do modelu zmiennej ukrytej, a więc takiej, która nie została zmierzona, ale ma wpływ przyczynowy na zmienne mierzalne występujące w modelu. Zasadniczy problem dotyczy niedoidentyfikowania parametrów, gdyż parametry modelu nie muszą być funkcją rozkładu zmiennych mierzalnych. Nawet jeśli parametry są identyfikowalne, to i tak rodzina rozkładów brzegowych dla zmiennych mierzalnych skojarzona z danym grafem nie posiada wielu własności, które ma rodzina rozkładów dla grafów bez zmiennych ukrytych (np. może nie istnieć MLE dla parametrów modelu).

Pearl (2000, 86-89) pokazuje, że nawet mimo nieidentyfikowalności parametrów danego modelu, możliwe jest zidentyfikowanie efektów manipulacji.

⁸⁸ Ten rodzaj przeszukiwania więzowego jest odmienny od przeszukiwania punktowego, gdzie poszczególnym modelom przypisuje się miejsce w szeregu. Ten drugi rodzaj przeszukiwania przestrzeni możliwych modeli przyczynowych jest wykorzystywany w przeszukiwaniu bayesowskim (Spirtes i in. 2004, 460-61).



Rysunek 4.13. (Opracowano na podstawie: Pearl 2000).

Stosując kryterium tylnych drzwi do grafu przedstawionego na Rysunku 4.13, otrzymujemy następujący efekt manipulacji na Z :

$$p(y | z \parallel p'(z)) = \sum_x p(y | x, z) p(x). \quad (4.42)$$

Kryterium tylnych drzwi pozwala wyznaczyć ten efekt, mimo nieidentyfikowalności parametrów modelu w tym przypadku.

W przypadku modeli ze zmiennymi ukrytymi nie można zastosować zarysowanej wcześniej strategii, gdyż tak zmodyfikowany algorytm byłby obliczeniowo niewykonalny,⁸⁹ a ponadto liczba grafów równoważnych rozkładowo mogłaby być nieskończona, jeśli nie wprowadziłoby się żadnych ograniczeń na zmienne ukryte wprowadzane do modelu.

W tym przypadku pojęcie równoważności rozkładowej zastępuje się pojęciem równoważności pod względem warunkowych niezależności dla zmiennych obserwowalnych. Dwa grafy należą więc do tej samej klasy, o ile wynikają z nich te same niezależności warunkowe dla zmiennych obserwowalnych. Analogiczne do pojęcia wzorca jest w tym przypadku pojęcie *częstkowego grafu poprzedników*. W takim grafie dwa węzły A oraz B sąsiadują ze sobą, gdy są zależne warunkowo względem każdego podzbioru zmiennych obserwowalnych $O \setminus \{A, B\}$. Skierowana krawędź $A \rightarrow B$ występuje w cząstkowym grafie poprzedników, jeśli A jest poprzednikiem B w każdym grafie należącym do klasy równoważności pod względem warunkowych niezależności dla O , a w żadnym grafie w tej samej klasie B nie jest poprzednikiem A .

⁸⁹ Takie przypadki W. Marciszewski proponuje określać w kategoriach „niedostępności algorytmicznej” (2004, 7).

Modyfikacją algorytmu PC, która dopuszcza zmienne ukryte, jest algorytm FCI. Zamiast wzorca, FCI określa cząstkowy graf poprzedników, dla którego jest punktowo zgodnym estymatorem.⁹⁰

Po zreferowaniu podstawowych dla podejścia BN zagadnień związanych z reprezentowaniem zależności przyczynowych oraz odkrywaniem ich na podstawie danych, powrócę do dyskutowanej w Punkcie 4.1 różnicy między *statystycznym* a *manipulacyjnym* podejściem do przyczynowości. Ujawnia się ona przede wszystkim na styku kwestii obserwacji i interwencji.⁹¹ Dobrej ilustracji dostarczają przykłady Pearla (2000, 155-59). Zacznę od podania przykładowych modeli przyczynowych w postaci równań strukturalnych. Niech model M_1 będzie dany przez następujące równania strukturalne:

$$x = u + \varepsilon_1 \quad (4.43)$$

$$z = \alpha x + \varepsilon_2 \quad (4.44)$$

$$y = \beta z + \gamma u + \varepsilon_3. \quad (4.45)$$

Natomiast model M_2 przez równania:

$$x = \varepsilon_1 \quad (4.46)$$

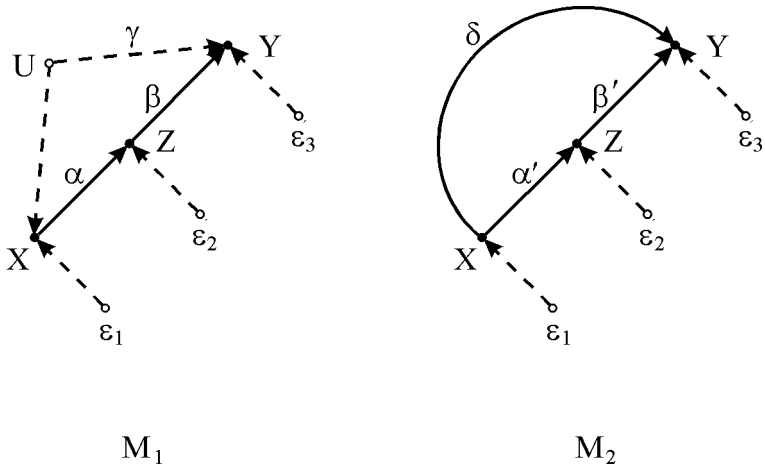
$$z = \alpha' x + \varepsilon_2 \quad (4.47)$$

$$y = \beta' z + \delta u + \varepsilon_3. \quad (4.48)$$

Rysunek 4.14 obrazuje graficzną różnicę pomiędzy tymi grafami dla tych modeli.

⁹⁰ Szczegółową dyskusję tego algorytmu podali M. Kłopotek i S. Wierchoń (2002, 227-52).

⁹¹ Mimo zbieżności terminologicznej, należy wyraźnie oddzielić parę: obserwacja-interwencja oraz badania eksperymentalne-nieeksperymentalne (obserwacyjne). Na podstawie modelu przyczynowego uzyskanego w badaniach nieeksperymentalnych można przewidywać efekt przyczynowy podjęcia określonego rodzaju działania. Więcej, podejście manipulacyjne traktuje kontekst rozważań związanych z podejmowanymi interwencjami jako pierwotny dla metodologicznego rozumienia przyczynowości. Dotyczy to w równym stopniu modeli przyczynowych w badaniach eksperymentalnych, jak i nieeksperymentalnych.



Rysunek 4.14. (Opracowano na podstawie: Pearl 2000).

Przewidywania oparte na obserwacji dotyczą sytuacji, gdy w oparciu o obserwację, że $X = x$, próbuje się podać odpowiedź na pytanie o efekt przyczynowy tej obserwacji na Y . Aby odpowiedzieć na takie pytanie, wystarczy znać rozkład łączny $P(X Y Z)$. Natomiast aby udzielić odpowiedzi na pytanie o efekt przyczynowy ustawienia X na wartość x poprzez podjęcie interwencji w danym modelu, trzeba określić zależności przyczynowe, gdyż inaczej nie da się przewidzieć, jak zmienią się zależności statystyczne. W notacji, którą wprowadził Pearl, różnicę tę oddają odpowiednio wyrażenia $E(Y | x)$ oraz $E(Y | do(x))$. Dla modelu M_1 otrzymujemy więc następujące zależności:

$$E(Y | x) = E(\beta\alpha x + \beta\epsilon_2 + \gamma x - \gamma\epsilon_1 + \epsilon_3) = (\alpha\beta + \gamma)x = r_{YX} \quad (4.49)$$

oraz

$$E(Y || x) = E(\beta\alpha x + \beta\epsilon_2 + \gamma u + \epsilon_3) = \alpha\beta x. \quad (4.50)$$

Druga z różnic między podejściem manipulacyjnym a statystycznym dotyczy przypadku, gdy równoważne obserwacyjnie dwa modele – w omawianym przykładzie są to M_1 oraz M_2 – można odróżnić na podstawie tego, że różnie przewidywać będą zmiany spowodowane daną interwencją.

Modele M_1 oraz M_2 są równoważne obserwacyjnie, gdyż przy ustawieniu $\alpha = \alpha'$, $\beta = \beta'$ oraz $\gamma = \delta$ oba modele dają te same przewidywania probabilistyczne dla zmiennych obserwowalnych Z , Y i Z (są względem tych zmien-

nych równoważne kowariancyjnie). Każdy jednak z modeli rozumianych jako „mechanizm generujący dane”, a nie jako skondensowany zapis danych statystycznych, daje inne przewidywania co do zmian, jakie w nich wystąpią wskutek zewnętrznej interwencji. I tak dla rozważanej wcześniej interwencji, która ustala wartość X na x , model M_1 daje następujące przewidywanie:

$$E(Y \parallel x) = E(\beta\alpha x + \beta\varepsilon_2 + \gamma u + \varepsilon_3) = \alpha\beta x. \quad (4.51)$$

Model M_2 natomiast

$$E(Y \parallel do(x)) = E(\beta'\alpha'x + \beta'\varepsilon_2 + \delta x + \varepsilon_3) = (\alpha'\beta' + \delta)x. \quad (4.52)$$

Wprowadzając wymogi dla równoważności kowariancyjnej w modelu M_2 , otrzymujemy zatem następujące przewidywanie:

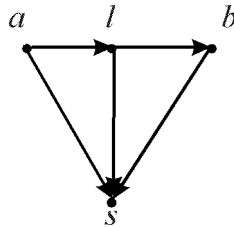
$$E(Y \parallel do(x)) = (\alpha\beta + \gamma)x. \quad (4.53)$$

Jednym z wyjaśnień tej różnicy może być podana przez Pearla interpretacja równań strukturalnych jako pewnego „stanu równowagi”, a nie jako niezmiennych równości matematycznych”. Zewnętrzne interwencje prowadzą do naruszenia tego stanu równowagi, a

moc modeli równań strukturalnych tkwi w tym, że rejestrują nie tylko stan równowagi wyjściowej, ale także informacje konieczne dla określenia, które równania muszą zostać naruszone, aby określić nowy stan równowagi. (2000, 157-58)

Można także przywołać znane z literatury przykłady takich modeli przyczynowych BN, które przy innych podejściach do modelowania zależności przyczynowych prowadzą do błędnego obliczenia efektów interwencji. Jeden z takich przykładów podał J. Robins (1997). Dotyczy on grupy pacjentów chorych na AIDS. Badane są cztery zjawiska: oddziaływanie zrandomizowane a , gdzie $X_a = 1$ oznacza, że pacjenta poddano działaniu AZT, a $X_a = 0$ oznacza, że pacjenta poddano działaniu placebo. Po pewnym okresie obserwowano wystąpienie u pacjentów zapalenia płuc l ($X_l = 1$). W odpowiedzi na tę sytuację pacjenci są poddawani leczeniu antybiotykami b . Przy czym wszystkich pacjentów chorych na zapalenie płuc poddaje się takiemu leczeniu: $P(X_b = 1 \mid X_l = 1) = 1$, u pozostałych natomiast to leczenie jest zrandomizowane. Po pewnym okresie obserwuje się, czy pacjent przeżył s ($X_s = 1$). Zasadnicze pytanie w tym badaniu dotyczy łącznego wpływu od-

działania leczenia AZT i antybiotyków na przeżycie: $P(X_s = 1 \mid X_a = 1, X_b = 1) = p(1_s \mid 1_a, 1_b)$.



Rysunek 4.15. (Opracowano no podstawie: Lauritzen 2001). Brak krawędzi $a \rightarrow b$ reprezentuje zrandomizowane przypisanie rodzaju oddziaływania.

W podejściu BN poszukiwany efekt przyczynowy jest dany wzorem (Lauritzen 2001, 80):

$$p(1_s \mid 1_a, 1_b) = \sum_{x_l} p(1_s \mid x_l, 1_a, 1_b) p(x_l \mid 1_a). \quad (4.54)$$

Zgodnie z tradycyjną analizą powinno się go skorygować względem l , które jest konfunderem w stosunku do efektu leczenia na umieralność (l bezpośrednio oddziałuje na s): $p(1_s \mid 1_a, 1_b) = \sum_{x_l} \hat{p}(1_s \mid x_l, 1_a, 1_b) \hat{p}(x_l)$.

Jednak l podlega oddziaływaniu leczenia i w związku z tym nie można skorygować l , lecz jedynie uwzględnić estymację prawdopodobieństwa warunkowego: $p(1_s \mid 1_a, 1_b) \sim \hat{p}(1_s \mid 1_a, 1_b)$. Oba rozwiązania dają wynik niezgodny z (4.54).

Jako podsumowanie różnic w interpretacji modeli przyczynowych między poszczególnymi podejściami do modelowania może służyć interpretacja strukturalnego charakteru parametrów na gruncie podejścia manipulacyjnego. Zgodnie z nią to, że równanie o postaci $y = \beta x + \varepsilon$ jest strukturalne, oznacza, że gdyby w idealnym eksperymencie ustawić zmienną X na wartość x oraz ustawić zmienne ze zbioru Z (nie zawierającego X ani Y) na wartość z , to wartość y zmiennej Y byłaby równa $\beta x + \varepsilon$, gdzie ε nie jest funkcją ustalenia x i z . (Pearl 2000, 160)

Z powyższej interpretacji wynika asymetryczność znaku równości występującego w równaniach strukturalnych, co odróżnia je od zwykłych równań algebraicznych. W przypadku obserwacji te równania są symetryczne,

jednak w przypadku interwencji są asymetryczne, gdyż nie określają, jak zachowa się zmienna X , gdy podjęta zostanie interwencja na Y .

Najmocniejsza empiryczna konsekwencja związana z taką interpretacją równania $y = \beta x + \varepsilon$ dotyczy wykluczenia innych zmiennych z prawej strony tego równania, gdyż wynika z tego testowalna empirycznie niezmienniczość – rozkłady zmiennej Y względem interwencji na X muszą pozostać niezmiennie przy manipulacji dowolną inną zmienną tego modelu. Symbolicznie: $P(y \parallel x \parallel z) = P(y \parallel x)$.

Dotyczy to jednak tylko interwencji, a nie obserwacji zmiennych, gdyż wartość $P(y \parallel x \mid z)$ będzie zależna od zaobserwowanej wartości z , jeśli pomiar ten dotyczy potomka zmiennej Y .⁹²

⁹² Własność niezmienniczości, o której tu mowa, nie dotyczy także zwykłego prawdopodobieństwa warunkowego $P(y \mid x)$, gdyż jego wartość będzie zmieniać się wraz z manipulowaniem innymi zmiennymi modelu (Pearl 2000, 161).

Kryterium proceduralne w odniesieniu do ilościowych modeli przyczynowych

Wspólnym odniesieniem dla dyskusji omówionych w poprzednim rozdziale na temat typów modeli przyczynowych w świetle kryterium proceduralnego jest, jak już pisałem, kwestia automatyzacji odkrywania zależności przyczynowych. Kluczową rolę odgrywa w tym przypadku, jak zostało to pokazane w Punkcie 4.4, tzw. założenie wierności, określane przez Pearla jako warunek stabilności. Wyklucza on przypadek, gdy zależność przyczynowa jest ściśle deterministyczna, gdyż wówczas nie byłoby podstaw do rozróżnienia między wieloma modelami danej zależności, a także przypadek, gdy zależność przyczynowa jest wsobnie losowa, gdyż nie istniałby jej model.¹ Tym samym założenie wierności wyklucza dwa najbardziej podstawowe ontologiczne rozumienia przyczynowości.²

A zatem, w konfrontacji z kryterium proceduralnym założenie wierności ujawnia swój przede wszystkim epistemologiczny charakter. Oparty na założeniu wierności program TETRAD realizuje, jak sędzę, odmienne od zamierzonego zadanie, a mianowicie zamiast ustalania zależności przyczynowych między badanymi zjawiskami określa zakres naszej (nie)wiedzy o tych zależnościach względem dostępnych informacji.

¹ A dokładniej, tzw. część systematyczna modelu; por. (Mynarski 2003).

² Por. (Krajewski 1967, 173-208).

5.1 Interpretacja przyczynowa modeli RCM wobec kryterium proceduralnego

Za cechę charakterystyczną podejścia RCM Rubin uznaje wyraźne oddzielenie reprezentacji zależności przyczynowych (ang. the science) od reprezentacji przypisania oddziaływania, czyli interakcji z badanymi zjawiskami. Dzięki temu – zdaniem Rubina i jego współpracowników – powstaje znacznie bardziej precyzyjna i informatywna reprezentacja zależności przyczynowych niż w przypadku modeli graficznych BN. Szczegółowe ilustracje tej różnicy Rubin zaprezentował podczas 19-tej Nordic Conference on Mathematical Statistics, która miała miejsce w 2002 r. w Sztokholmie (Palmgren 2004), a którą opublikowano w *Scandinavian Journal of Statistics* w 2004 r.³

Pierwszy, hipotetyczny przykład podany przez Rubina dotyczy różnicy między skutecznością standardowej metody ($W = 0$) a nowej metody operacyjnej ($W = 1$). Uwzględniano zmienne niepodlegające oddziaływaniu (X , np. wiek pacjenta) oraz zmienne kontrfaktyczne, reprezentujące skuteczność metody operacyjnej: $Y(0)$ – ilość lat przeżytych po operacji nowej oraz $Y(1)$ – standardową metodą. W celu uniknięcia nieporozumień związanych z podejściem RCM, Rubin podkreśla, że w momencie poprzedzającym faktyczne przeprowadzenie operacji obydwie zmienne *nie* są kontrfaktyczne, gdyż możliwe jest pobranie ich wartości dla dowolnego z pacjentów. Jednak po przeprowadzeniu operacji ($W = 0, 1$) możemy pobrać faktyczną wartość tylko jednej z dwóch zmiennych – natomiast wartość drugiej jest wartością „brakującą”. Przy założeniu, że nie można ocenić zachodzenia zależności przyczynowych *simpliciter*, lecz tylko porównawczo (Rubin 1990a, 280), ustalanie zależności przyczynowych sprowadza się do wnioskowania z brakującymi danymi: dla danej jednostki posiadamy tylko jedną z dwóch (lub więcej – przy większej liczbie uwzględnianych oddziaływań) potrzebnych wartości.

Posługując się założeniem SUTVA, zgodnie z którym oddziaływanie na jednostkę nie zmienia się wraz z interwencją na inne jednostki (Rubin 1980; 1990b), oblicza się brakujące wartości dla zmiennych kontrfaktycznych. Średnia zaobserwowana wartość $Y(0)$ w podanym przykładzie wynosi 5,4, a rzeczywista (a więc obliczona z uwzględnieniem wartości kontrfaktycznych niezaobserwowanych) 7. Dla $Y(1)$ odpowiednio: 5 oraz 11. Ponadto, zaob-

³ (Aalen 2004; Arjas i Parnier 2004a; 2004b; Lauritzen 2004; Palmgren 2004; Rubin 2004a, 2004b).

serwowane wartości $Y(0)$ mieszczą się w przedziale (4, 6), a dla $Y(1)$ w przedziale (9, 14). Natomiast wszystkie (a więc łącznie z kontrfaktycznymi) wartości $Y(0)$ zawarte są w przedziale (4, 13), a dla $Y(1)$ w przedziale (0, 14). Uwzględnienie tylko danych zaobserwowanych może nasunąć wniosek, że – biorąc pod uwagę efekt średni – skuteczniejsza jest nowa metoda, gdyż najniższa zaobserwowana wartość $Y(1)$ jest i tak wyższa od najwyższej zaobserwowanej wartości $Y(0)$. Taki wniosek byłby uzasadniony, gdyby dodatkowo słuszne było założenie, iż dobór metody operacyjnej dla poszczególnych pacjentów był losowy (5 osób poddanych standardowej metodzie i 3 osoby w grupie kontrolnej):

$$P(W | X, Y(0), Y(1)) = 1/56, \text{ jeśli } \sum W_i = 3, \quad (5.1)$$

a 0 w przeciwnym wypadku.

Jednakże w tym przypadku mechanizm przypisania oddziaływania, a więc wybrania rodzaju operacji dla danego pacjenta, jest zależny od tych wartości zmiennych kontrfaktycznych tak, aby uzyskać możliwie najlepszy efekt leczenia dla danego pacjenta, jak wyraża to formuła:

$$P(W | X, Y(0), Y(1)) = \prod_{i=1}^8 P(W_i = 1 | X_i, Y_i(0), Y_i(1)), \quad (5.2)$$

gdzie $P(W_i = 1 | X_i, Y_i(0), Y_i(1)) = 1$, jeśli $Y_i(0) \geq Y_i(1)$, a 0 w przeciwnym wypadku.

Oceny skuteczności dwóch metod operacyjnych w podejściu RCM dokonuje się na podstawie różnicy średnich wartości tych dwóch zmiennych (dla wszystkich pacjentów): $Y_i(1) - Y_i(0)$. W analizowanym przykładzie wynosi ona -2 , co jednoznacznie wskazuje, że skuteczniejsza jest operacja standardowa.

Jak ocenić ten przykład w świetle kryterium proceduralnego? Zgodnie z definicją oddziaływania czynnego, w przypadku 5 osób poddanych operacji standardowej ta operacja oddziaływała na nich czynnie, a w przypadku 3 osób poddanych operacji nową metodą była ona dla nich oddziaływaniem czynnym. Uwzględniając współzmienną X (wiek pacjentów), otrzymujemy następujące zbiory danych interwencyjnych:

X	W	Y
66	0	4
68	1	14
70	0	6
72	0	6
72	1	9
76	0	6
81	0	5
81	1	10

a)

X	W	Y
66	1	1
68	0	13
70	1	3
72	1	1
72	0	8
76	1	0
81	1	2
81	0	8

b)

Tabela 5.1. Dane: a) interwencyjne w przypadku populacji 8 osób poddanych próbie medycznej; b) w przypadku większej populacji, z której wybrano te 8 osób.

Rubin łączy te dwa zbiory danych interwencyjnych,⁴ pochodzących z różnych interwencji dla różnych populacji, za pomocą założenia SUTVA. Dzięki temu otrzymuje, w sposób opisany wcześniej, jasną odpowiedź na pytanie o to, który z rodzajów operacji jest skuteczniejszy.

Jeśli jednak uwzględnić proponowane tu podejście proceduralne, to okazuje się, że mamy do czynienia z pewną trudnością, o której Rubin nie wspomina w swojej analizie. Dobrane przez niego dane, zaprezentowane w Tabeli 5.1, stanowią dwa odrębne zbiory danych interwencyjnych. Jeden dla badanych 8 pacjentów, a drugi dla szerszej populacji, z której oni pochodzą. Jeśli przyjąć wcześniejsze oznaczenia, to oczywiście są w tym przypadku zależności $W \rightarrow Y$ oraz $X \rightarrow Y$. Pozostaje pytanie o zależność $X \rightarrow W$. Dla danych 8 pacjentów z kolumny a) w Tabeli 5.1 okazuje się, że warunkowanie względem Y nie usuwa zależności między W a X . Podobnie jest, gdy – stosując założenie SUTVA – połączymy dane z kolumn a) i b). Analiza samych danych z kolumny b) dla ogólnej populacji prowadzi natomiast do wniosku, że W i X są warunkowo niezależne względem Y . Oprócz tego sama analiza danych przeprowadzona osobno dla kolumn a) i b) prowadzi do niezgodnego z wiedzą ogólną wniosku, że W i Y są niezależne.

Ponadto zasadnicza dla tego przykładu zależność $W \rightarrow Y$, która występuje we wszystkich trzech przypadkach, ma współczynnik pozytywny dla

⁴ Nie jest jasny status danych dla populacji ogólnej. Na podstawie samego założenia SUTVA nie można traktować ich jako danych interwencyjnych dla 8 pacjentów opisanych w kolumnie a) Tabeli 5.1. Czy natomiast jest to zbiór danych interwencyjnych w przyjętym tu sensie dla ogólnej populacji, nie może być jednoznacznie stwierdzone na podstawie opisu podanego w (Rubin 2004). Przyjmuję jednak, iż tak jest.

populacji 8 pacjentów, a negatywny dla populacji ogólnej oraz połączonych populacji.

Niezgodności wykazane przy stosowaniu kryterium proceduralnego w odniesieniu do dwóch zbiorów (a w zasadzie trzech, biorąc pod uwagę ich sumę) danych analizowanych przez Rubina nasuwa wniosek, że w tym przypadku nie można jednoznacznie stwierdzić zależności przyczynowej w badanej dziedzinie.

Przyjęte w tej pracy kryterium proceduralne prowadzi do wniosku, że aby móc orzec, iż jest to model zależności przyczynowych w badanej dziedzinie, należy uzyskać dodatkowe dane interwencyjne. Jednym z możliwych źródeł jest określenie pola ekspansji dla kluczowej w tym przypadku zależności $W \rightarrow Y$ i uzyskanie informacji o dodatkowych czynnikach, które powodują, że tylko w przypadku 8 pacjentów jest to zależność pozytywna.

Wobec powyższych uwag nasuwa się wniosek, że być może w przedstawionym przez Rubina przykładzie mamy do czynienia z przyczynowością względem modelu. Zwróćmy bowiem uwagę na fakt, że założenie SUTVA, dzięki któremu Rubin łączy te dwa zbiory danych w jeden i następnie określa znak zależności między metodą operacji W a ilością przeżytych później lat Y , nie wystarcza do tego, by potraktować ten nowy zbiór danych jako dane interwencyjne. Dane z populacji ogólnej Rubin traktuje jako uzupełnienie brakujących danych o 8 pacjentach, np. dla pacjenta operowanego nową metodą oblicza w ten sposób – kontrfaktyczną – długość życia, gdyby był operowany starą metodą. Następnie porównuje te dwie liczby, a średnia różnica wynosząca -2 dla wszystkich 8 osób daje podstawę do tego, by uznać metodę standardową za lepszą. Tak uzupełnione dane kontrfaktyczne nie spełniają jednak warunku, o którym mowa w definicji danych interwencyjnych – brak jest bowiem faktycznie *zaobserwowanej* odpowiedzi danego pacjenta na oddziaływanie, którym w tym przypadku jest wybór określonej metody operowania.

Z tego więc powodu nie stosuje się do tego przykładu kryterium proceduralnego zależności przyczynowej w badanej dziedzinie, lecz wyłącznie względem modelu. Przyjmując za Rubinem założenie SUTVA, można poddać analizie powstały z połączenia dwóch zbiorów danych, a wyniki analizy uznać – przy spełnieniu kryterium proceduralnego dla modelu – za ustalenie zależności przyczynowej względem modelu.

Za zasadnicze w tym przypadku składowe modelu należy uznać:

- 1) założenie o przyczynowej wystarczalności zmiennych X , Y oraz W ;

- 2) założenie SUTVA;
- 3) założenia wynikające z planu badań i wiedzy ogólnej: odrzucenie zależności $Y \rightarrow X$ oraz $Y \rightarrow W$, a także $W \rightarrow X$.

Założenie 1. można skrótowo wyrazić jako tezę, że dla konstruowanego modelu ze zbiorem zmiennych V nie istnieje wspólna przyczyna dla żadnej pary zmiennych w V , która nie należałaby do tego zbioru. To założenie jest jednak wątpliwe, jak zwróciłem już na to uwagę. Zasadniczym powodem są rozbieżności zaobserwowane w populacji ogólnej oraz podpopulacji, co sugeruje istnienie ukrytej zmiennej, która nie jest reprezentowana w modelu, a która zmienia znak zależności $W \rightarrow Y$. Jednym z możliwych sposobów uchylecia tej trudności jest skonstruowanie modelu, który nie przyjmuje założenia 1., a więc modelu przyczynowego ze zmiennymi ukrytymi.

5.2 Rozbieżności między typami modeli przyczynowych wobec kryterium proceduralnego

Założenia wyliczone w Punkcie 5.1 są milcząco przyjęte przez Rubina. W przypadku analizy danych przez program TETRAD ich jawna reprezentacja jest jednak konieczna dla uzyskania poprawnego modelu. W przeciwnym wypadku program podałby wzorzec z niezorowanymi ścieżkami.

Zanim przejdę do omówienia analizy tego przykładu za pomocą programu TETRAD, skupię uwagę na kolejnym z przykładów podanych przez Rubina, który miał wykazać „wyższość” podejścia RCM w stosunku do BN. Jest to uproszczona wersja testów przeprowadzanych na zlecenie Centers for Disease Control and Prevention, mających określić dawkę szczepionki skutecznej w przypadku narażenia na śmiertelną dawkę węgla oraz bezpieczną dawkę tej szczepionki, która minimalizuje skutki uboczne. Przeprowadzono dwa testy: w pierwszym uczestniczyli wolontariusze, których w trakcie testu poddawano licznym pomiarom reaktogeniczności (wskaźnik skutków ubocznych) oraz odporności, w drugim natomiast uczestniczyły makaki, poddane analogicznym pomiarom, ale dodatkowo mierzono w ich przypadku przeżycie po podaniu dawki 500 razy większej niż śmiertelna dla nieszczepionych osobników. W tym przypadku zasadnicza kwestia dotyczyła tego, jak efekty poszczególnych dawek szczepionki zaobserwowane u makaków, odnieść do przypadku ludzi. Kluczem do rozwiązania było potraktowanie pomiarów odporności u ludzi jako biomarkerów lub rezultatów zastępczych przeżycia i wywnioskowanie szans przeżycia na podstawie stwierdzonej u makaków relacji: poziom odporności-przeżycie.

Rubin przyjmuje dwa poziomy szczepionki: $W = 1$ (niski) oraz $W = 2$ (wysoki) u makaków i analogicznie u ludzi: $W^* = 1$ oraz $W^* = 2$. Poziom odporności P oraz P^* to wielkość skalarna (złożone zjawisko reprezentuje się za pomocą jednej liczby). Przeżycie przy poddaniu oddziaływaniu jednym z poziomów szczepionki to zmienna Y , przyjmująca dwie wartości: 1 (żywy) oraz 0 (nieżywy). Wartości dla Y^* są niezaobserwowane w tym teście. W swojej analizie Rubin wprowadza zmienne kontrfaktyczne: $P(1) = L$, H to poziom odporności danej jednostki po zastosowaniu dawki $W = 1$, podobnie dla $P(2)$ oraz $Y(1)$ i $Y(2)$ – zmienne reprezentujące przeżycie jednostki po zastosowaniu dawki niskiej lub wysokiej.

Zasadniczy problem dotyczy określenia poziomu szczepionki, który chroni przed chorobą, a więc wartości $Y^*(1)$ przy oddziaływaniu niskiego poziomu oraz $Y^*(2)$ przy oddziaływaniu poziomu wysokiego. Zasadniczym krokiem w znalezieniu rozwiązania jest przyjęcie, że $W = W^*$ oraz $P = P^*$ i na podstawie danych dostępnych dla makaków dokonanie przewidywania wartości $Y^*(1)$ oraz $Y^*(2)$. Jednym z zasadniczych warunków poprawności takiego przewidywania jest, aby oddziaływanie W na Y było w całości zapośredniczone przez P (podobnie dla ludzi). Właśnie w tym Rubin (2004, 166) dostrzega trudność podejścia graficznego, które miałyby się ograniczyć wyłącznie do zaobserwowanych wartości zmiennych W , P oraz Y . Trudność tę ilustrują hipotetyczne dane z Tabeli 5.2.

$P_{obs} = L$		$P_{obs} = H$	
W_{obs}	Y_{obs}	W_{obs}	Y_{obs}
1	20	1	80
1	20	2	80
2	20	2	80

a)

$P_{obs} = L$		$P_{obs} = H$	
W_{obs}	Y_{obs}	W_{obs}	Y_{obs}
1	20	1	80
1	20	2	70
2	0	2	70

b)

Tabela 5.2. Dane zaobserwowane dla testu szczepionki na węgliku: a) W_{obs} oraz Y_{obs} są niezależne warunkowo względem P_{obs} ; b) W_{obs} oraz Y_{obs} są zależne warunkowo względem P_{obs} (Y_{obs} oznacza proporcję osobników, które przeżyły). (Opracowano na podstawie: Rubin 2004).

Dla danych w kolumnie a) W_{obs} oraz Y_{obs} są warunkowo niezależne. Dla poszczególnych kategorii P_{obs} zaobserwowana wartość Y_{obs} nie zmienia się wraz ze zmianą W_{obs} . Przeciwnie jest natomiast dla danych w kolumnie b), gdzie dla poszczególnych kategorii P_{obs} zmianie wartości W_{obs} towarzyszy zmiana Y_{obs} .

Problem metod graficznych, twierdzi Rubin, polega na tym, że w analizie samych zaobserwowanych danych nie można osiągnąć właściwego wniosku, do którego prowadzi podejście RCM. W przypadku a), gdzie zaobserwowane dane prowadzą do stwierdzenia warunkowej niezależności W i Y , analiza RCM ustala, że W ma bezpośredni wpływ przyczynowy na Y warunkowo względem $P(1)$ oraz $P(2)$. W przypadku b) natomiast przeciwnie: mimo warunkowej zależności W i Y na podstawie danych obserwacyjnych analiza kontrfaktyczna ustala, że W ma wyłącznie pośredni wpływ przyczynowy na Y warunkowo względem $P(1)$ oraz $P(2)$.

Analiza Rubina w kategoriach zmiennych kontrfaktycznych prowadzi do rozbicia danych przedstawionych w Tabeli 5.2 wedle wyróżnionych zasadniczych stratyfikacji (por. Punkt 4.2). W analizowanym przykładzie są to trzy grupy: o niskiej odporności, dla której poziom odporności nie zmienia się wraz ze zwiększeniem dawki szczepionki $P(1) = P(2) = L$, o normalnej odporności, gdzie zwiększenie dawki zwiększa odporność $P(1) = L \neq P(2) = H$ oraz o wysokiej odporności, niezależnej od dawki szczepionki $P(1) = P(2) = H$. Nową postać danych przedstawia Tabela 5.3.

	$W_{obs} = 1$			$W_{obs} = 2$		
	$P(1), P_{obs}$	$Y(1)$	\bar{Y}_{obs}	$P(2), P_{obs}$	$Y(2)$	\bar{Y}_{obs}
a)	L	0	20	L	20	20
	L	40		H	60	
	H	80		H	100	
b)	L	0	20	L	0	0
	L	40		H	60	
	H	80		H	80	

Tabela 5.3. Dane dla testu szczepionki węglika w analizie kontrfaktycznej: a) bezpośredni wpływ W na Y warunkowy względem $S(1)$ i $S(2)$ przy niezależności warunkowej dla zaobserwowanych wartości tych zmiennych; b) brak bezpośredniego wpływu W na Y warunkowego względem $S(1)$ i $S(2)$ przy zależności warunkowej dla zaobserwowanych wartości tych zmiennych. (Opracowano na podstawie: Rubin 2004).

Tabela 5.3 zachowuje trzy zasadnicze stratyfikacje danych przyjęte przez Rubina, czyli LL, LH oraz HH, zamiast jednak uwyrażniać efekt kontrfaktyczny, jak ma to miejsce w oryginale (Rubin 2004, 166), przedstawiłem ją w sposób pozwalający łatwiej odczytać przejście od danych w postaci rezultatów potencjalnych z przyjętą przez Rubina stratyfikacją do danych zaobserwowanych, przedstawionych w Tabeli 5.2. W panelu a) w przypadku, gdy $W = 1$, wyróżniają się dwie grupy pod względem poziomu odporności:

$P_{\text{obs}} = L$ oraz $P_{\text{obs}} = H$. Ta pierwsza obejmuje jednak dwie zasadnicze stratyfikacje: LL oraz LH o zasadniczo odmiennych proporcjach przeżycia: $Y_{LL}(1) = 0\%$ oraz $Y_{LH}(1) = 40\%$, czego średnią jest zaobserwowana wartość $\bar{Y}_{\text{obs}} = 20\%$. Natomiast gdy $W = 2$, wówczas zaobserwowana grupa $P_{\text{obs}} = H$ ulega podziałowi na dwie zasadnicze stratyfikacje: LH oraz HH o proporcjach przeżycia: $Y_{LH}(2) = 60\%$ oraz $Y_{HH}(2) = 100\%$, czego średnią jest zaobserwowana wartość $\bar{Y}_{\text{obs}} = 80\%$.

W drugim przypadku, który przedstawiony jest w panelu b), otrzymujemy te same wartości co w panelu a) dla przypadku $W = 1$. Różnica daje się zauważyć, gdy $W = 2$. Wówczas $Y_{LL}(2) = 0\%$, $Y_{LH}(2) = 60\%$ oraz $Y_{HH}(2) = 80\%$, czego średnią jest zaobserwowana wartość $\bar{Y}_{\text{obs}} = 70\%$ dla obu stratyfikacji w kategorii $P_{\text{obs}} = H$.

Z punktu widzenia analizy RCM najistotniejsza różnica w obu przypadkach (wyróżnionych w panelu a) oraz b)) sprowadza się do tego, jak zachowuje się różnica $Y(1) - Y(2)$ w poszczególnych zasadniczych stratyfikacjach (LL, LH oraz HH). W panelu a) jest ona stała, z czego Rubin wnioskuje o bezpośrednim wpływie W na Y . W panelu b) tę różnicę można zaobserwować wyłącznie w jednej zasadniczej stratyfikacji, a mianowicie LH, czyli takiej, gdzie W zmienia też poziom odporności organizmu. W tym więc przypadku Rubin wyciąga wniosek, że wpływ przyczynowy W na Y jest w całości zapośredniczony przez P , czyli poziom odporności organizmu – jeśli W nie zmienia tegoż poziomu, to nie zmienia także przeżycia w danej grupie.

Łącznie oba przykłady z panelu a) oraz b) pokazywałyby, że warunkowa (nie)zależność zmiennych stwierdzona na podstawie analizy danych pochodzących z obserwacji nie jest warunkiem koniecznym ani wystarczającym odpowiedzi na pytanie, czy $X \rightarrow Y$ jest zależnością przyczynową, czy też nie.

Zwróćmy uwagę, że z punktu widzenia podejścia proceduralnego mamy w tym przypadku do czynienia z dwoma zbiorami danych interwencyjnych, jak przedstawia to panel a) oraz b) Tabeli 5.2. Nie tylko analiza RCM, ale – na co zwraca uwagę Rubin – również analiza BN, oparta wyłącznie na danych zaobserwowanych, prowadzi do wniosków odmiennych dla panelu a) oraz b).

Dla Rubina istotne jest jednak to, że w przypadku obu rodzajów analizy wnioski są przeciwstawne. Ten element analizy jest w tym miejscu mniej istotny, jak wyjaśnię to w dalszej części tego rozdziału, omawiając komentarz Lauritzena do tekstu Rubina. Ważne natomiast jest zachodzenie

tych niezgodności, gdyż wskazuje ono, że – podobnie jak miało to miejsce w poprzednio omawianym przykładzie – kryterium proceduralne nie jest w tym przypadku poprawnie zastosowane. W zastosowaniu tego kryterium nie uwzględniono bowiem zjawiska – lub ich kompleksu – które sprawia, że między dwoma próbkami (panel a) oraz b)) obserwujemy tak znaczącą różnicę. Podobnie więc jak w przykładzie wcześniejszym, Rubin operuje danymi z dwóch próbek (populacji), które istotnie są od siebie różne, i danych nie można uznać za kompletne.

Ponownie więc założenie 1., określone wyżej jako założenie przyczynowej wystarczalności, okazuje się nie stosować do omawianej sytuacji. Do kwestii, jak na gruncie BN uzyskać poprawny model dla jednej z dwóch sytuacji z panelu a) lub b), powrócę w dalszej części rozdziału.

Punkt ciężkości omawianego przykładu tkwi w roli poziomu odporności jako surogatu/biomarkera przeżycia jednostki. Przede wszystkim, ważna jest odpowiedź na pytanie, jak przenieść wyniki otrzymane w badaniach na makakach na ludzi. Z proceduralnego punktu widzenia jest to pytanie o to, czy testy przeprowadzane na makakach są *ekspansją* pola interwencji dla relacji między dawką szczepionki u ludzi W^* a ich poziomem odporności P^* . Rubin proponuje następujące rozwiązanie:

Nadzieja w tym, że P^* oraz P „mają to samo znaczenie” dla przeżycia w tym sensie, że P jest „surogatem/biomarkerem” przeżycia u makaków poddanych szczepieniu i podobnie P^* jest „surogatem/biomarkerem” u ludzi tak, że nie ma „bezpośredniego” efektu W na Y warunkowo względem P , ani W^* na Y^* warunkowo względem P^* . Przypadkiem, gdzie można szukać uzasadnienia dla tego twierdzenia, są makaki, dla których zmierzono wszystkie wielkości: W , $S(W)$ oraz $Y(W)$. (Rubin 2004, 165)

W odpowiedzi na to pytanie Rubin podaje przykłady ujęte w Tabeli 5.3. Ich celem jest przede wszystkim ukazanie, że warunkowa (nie)zależność zmiennych określana na podstawie danych z obserwacji nie jest ani warunkiem wystarczającym, ani koniecznym stwierdzenia, czy W ma bezpośredni wpływ na Y warunkowo względem P . Do tej kwestii powrócę poniżej, dyskutując analizę tego przykładu z perspektywy podejścia graficznego BN. Teraz uwagę skupimy na problemie, czy przejście od zjawisk W , Y oraz P do W^* , Y^* oraz P^* jest *ekspansją* pola interwencji $W^* \rightarrow P^*$.

W przyjętym przeze mnie kryterium proceduralnym *ekspansją* pola interwencji dokonuje się ze względu na dekompozycję w oparciu o relację bycia maksymalnie zwartą częścią fuzji mereologicznej (por. Punkt 2.2). Oczywiście w omawianym przykładzie nie mamy bezpośrednio do czynienia

z taką relacją. Ale czy *funkcjonalnie*, jak sugeruje to Rubin, a w komentarzu także Lauritzen, nie mamy do czynienia ze zjawiskami podobnymi? Czy z perspektywy funkcjonalnej nie można wyodrębnić części wyznaczających jednoznacznie poszukiwaną dekompozycję? Spróbuję rozwinąć tę sugestię.

W przypadku ludzi analizowany przykład testu polega na zaaplikowaniu dawki szczepionki na węglika o różnym stężeniu (L lub H). Obserwujemy następnie zmiany w poziomie odporności. Brak jest natomiast obserwacji przeżycia jednostek o różnym poziomie odporności po poddaniu oddziaływaniu 500 razy większej dawki niż śmiertelna. Takie obserwacje są dostępne tylko dla makaków. Można przyjąć, jak proponuje Rubin, że poziom odporności jest surogatem/biomarkerem, i sprawdzić tę hipotezę w przypadku makaków, stwierdzając, czy wpływ W na Y jest w całości zapośredniczony przez P . Potwierdzenie takiej hipotezy oznaczałoby, że można przyjąć, iż P^* jako funkcja P byłaby podstawą przewidywania Y^* dla ludzi. Zwróćmy uwagę, jak Rubin, że poziom odporności nie jest prostym zjawiskiem. W tym przypadku szczególnie istotną rolę odgrywa jego rozłożenie w czasie. W poziomie odporności ludzi oraz makaków wyróżniałyby się naturalnie dwie części: przed podaniem i po podaniu szczepionki. O funkcjonalnej odpowiedniości można byłoby mówić tylko w odniesieniu do tej drugiej części złożonego zjawiska, którym jest poziom odporności. Zatem P reprezentujące tę właśnie część miałyby funkcjonalny odpowiednik w postaci P^* . Tym samym P , poprzez tę funkcjonalną odpowiedność, stawałoby się częścią poziomu odporności. Stąd nasuwałby się wniosek, że mamy do czynienia z ekspansją pola interwencji.

Droga do ustalenia tego wniosku jest jednak znacznie bardziej złożona. W odniesieniu do szczepionki na węglika dane J. Wang i M. Roehrl (2005) przekonują, że jej skuteczność jest bardzo zróżnicowana dla poszczególnych gatunków ssaków. Może wynikać to z faktu, że nie do końca jasna jest rola poszczególnych składników szczepionek w podnoszeniu poziomu odporności. Ponadto w przypadku zwierząt szczepionka weterynaryjna jest znacznie bardziej skuteczna niż stosowana u ludzi, mimo tego, że „często wzbudza niższy poziom przeciwciał” (Wang i Roehrl 2005, 4).

D. A. Freedman i H. Zeisel (1988a; 1988b) zwracają także uwagę na brak systematycznych podstaw do automatycznego przenoszenia wyników testów zaobserwowanych u zwierząt na ludzi. Jeden z głównych problemów to różnica dawki, która w badaniach laboratoryjnych na zwierzętach może być duża, co znacznie odbiega od – niekiedy śladowych – obecności w naturalnym środowisku człowieka.

Ta dyskusja prowadzi do wniosku, że odpowiedniość funkcjonalną między P a P^* można jedynie uznać na podstawie przyjętego w modelu założenia. Nie ma ona pełnego ugruntowania teoretycznego, a wyniki eksperymentalne otrzymane w odniesieniu do badań szczepionki na wąglika wykazują duże zróżnicowanie międzygatunkowe. Dlatego więc w tym przypadku nie można uznać relacji $P \rightarrow Y$ za ekspansję pola interwencji $W^* \rightarrow P^*$. Takie przejście jest uzależnione od przyjętego w modelu założenia o rodzaju odpowiedności między P^* a P . Przytoczone powyżej racje teoretyczne i eksperymentalne poddają jednak w wątpliwość założenie o funkcjonalnej odpowiedności surogatów P i P^* .

Ogólny wniosek z analizy tego przykładu jest więc następujący: jeśli uznać zależność wskazaną przez Rubina w analizie RCM za przyczynową, to zgodnie z kryterium podanym w Punkcie 2.3 jest to co najwyżej zależność przyczynowa względem modelu. Kryterium proceduralne dla tego przypadku jest bowiem spełnione tylko pod warunkiem przyjęcia założeń:

- 1) założenie o przyczynowej wystarczalności zmiennych W , Y oraz Y i W^* , P^* oraz Y^* ;
- 2) założenie o funkcjonalnej odpowiedności surogatów/biomarkerów P i P^* ;
- 3) założenia wynikające z planu badań i wiedzy ogólnej.

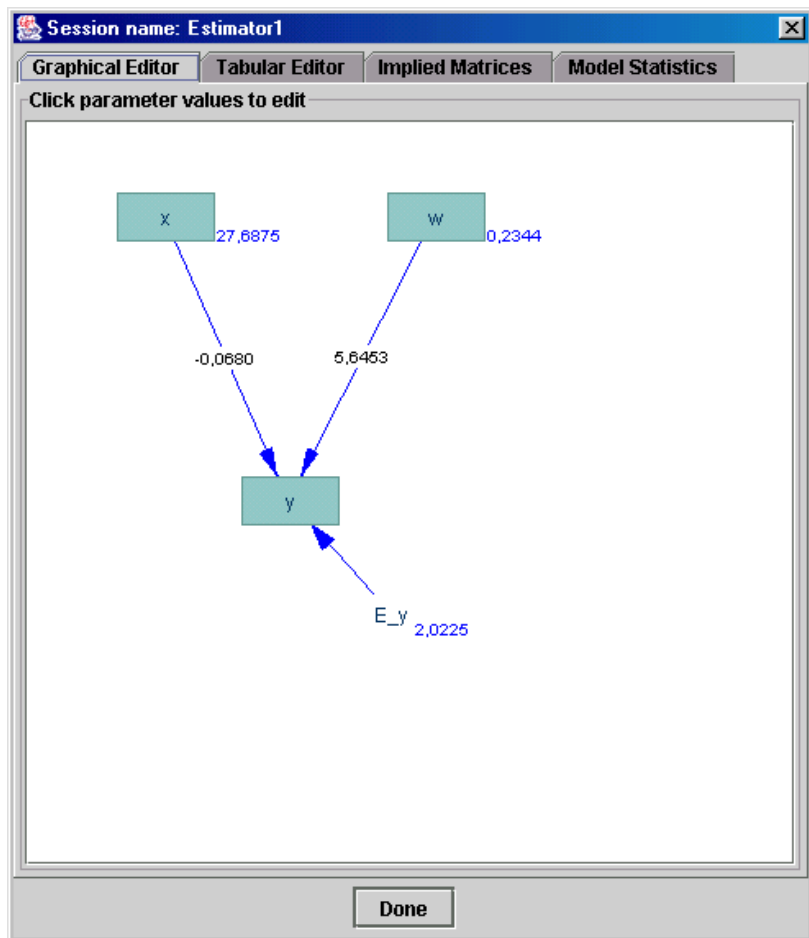
Wymienione wcześniej zastrzeżenia poddają w wątpliwość dwa pierwsze założenia. Różnica między danymi w panelu a) i b) Tabeli 5.3 nasuwa wniosek, że między tymi dwoma obserwowanymi populacjami zachodzi różnica, która nie jest jawnie reprezentowana w modelu.

Po analizie obu przykładów z perspektywy podejścia RCM, przejdźmy do dyskusji na ich temat na gruncie modelowania graficznego BN.

5.3 Interpretacja przyczynowa modeli BN wobec kryterium proceduralnego

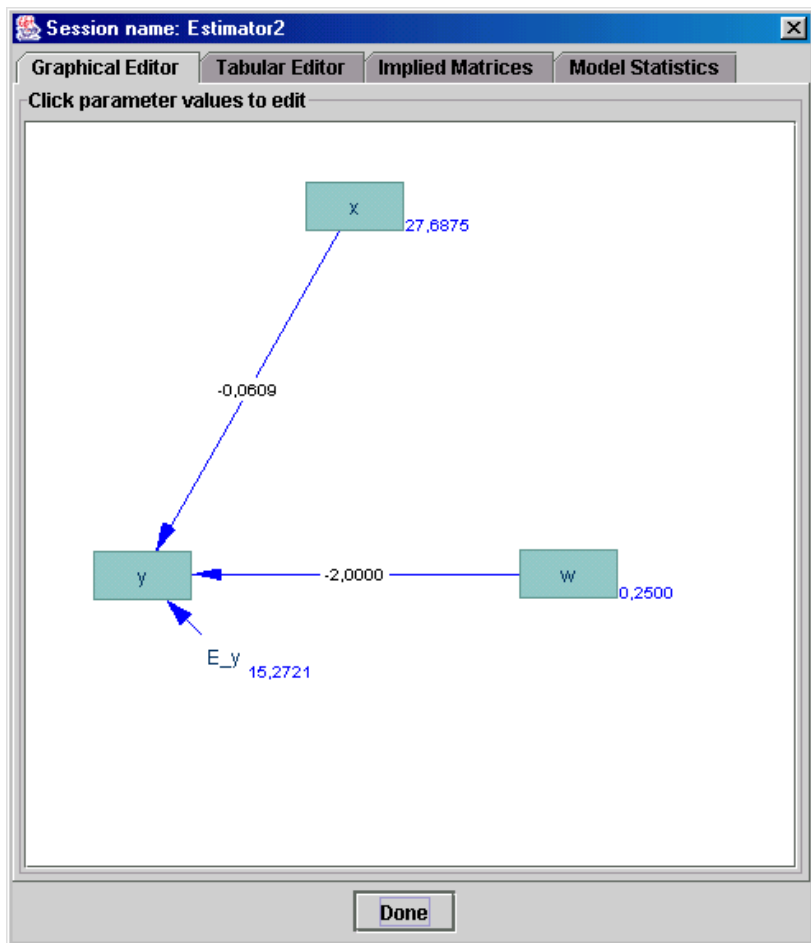
Na początku przywołam omawiany w Punkcie 5.1 przykład Rubina dotyczący skuteczności standardowej i nowej metody operacyjnej. Poniżej przedstawiłem wyniki analizy danych dla tego przykładu za pomocą programu TETRAD. Dane pochodzą z Tabeli 5.1, a przy analizie korzystałem z założeń przyjętych przez Rubina. Rysunek 5.1 przedstawia model graficzny, uzyskany przez zastosowanie algorytmu PC (założenie o przyczynowej wystarczalności), wraz z estymowanymi parametrami. Zasadnicza różnica w stosunku do modelu dla populacji ogólnej dotyczy znaku zależności $W \rightarrow Y$.

Zastosowanie algorytmu PC programu TETRAD do danych hipotetycznych Rubina (Rubin 2004) prowadzi do podobnego wniosku, jak analiza RCM: nowa metoda operacyjna jest mniej skuteczna od standardowej. Obliczony przez TETRAD współczynnik tej zależności ma dokładnie tę samą wartość, co w przypadku analizy kontrfaktycznej przy podejściu RCM. Graficzną reprezentację modelu odkrytego z danych przez TETRAD podaje Rysunek 5.2.



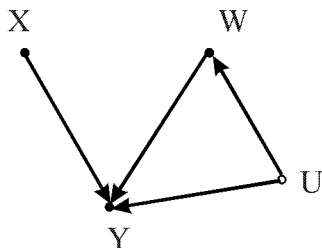
Rysunek 5.1. Model skuteczności nowej metody operacyjnej w populacji 8 pacjentów przy założeniu wystarczalności przyczynowej.

W związku z rozbieżnościami, które zaobserwowano w stosunku do obu tych przypadków, konieczne jest, jak podkreślałem w Punkcie 4.2, odrzucenie założenia o wystarczalności przyczynowej. W tym przypadku bowiem należy przyjąć istnienie ukrytej zmiennej, np. U , która konfunduje relację między W a Y . Jednym ze standardowych sposobów uniknięcia tego rodzaju konfundowania jest ustalenie takiego planu badań, w którym stosowanie metody operacyjnej jest przypisywane pacjentom losowo. W omawianym przykładzie jednak taki plan badawczy nie był zastosowany.



Rysunek 5.2. Model skuteczności nowej metody operacyjnej w populacji ogólnej przy założeniu wystarczalności przyczynowej.

Program TETRAD napotyka na zasadniczą trudność przy konstruowaniu grafu przedstawionego na Rysunku 5.3. W omawianym przypadku dane pochodzą z populacji ogólnej oraz jej podpopulacji. A ponieważ w obu przypadkach zachodzi rozbieżność rezultatów, brakuje więc rozstrzygnięcia, który ze zbiorów danych jest istotniejszy. Zastosowanie jednego z algorytmów, jak FCI, do poszukiwania modelu przyczynowego z dopuszczeniem istnienia zmiennych ukrytych nie tylko nie prowadzi, ale i nie może prowadzić do odkrycia grafu przedstawionego na Rysunku 5.3.



Rysunek 5.3. Model skuteczności nowej metody operacyjnej (gdzie U jest zmienną ukrytą).

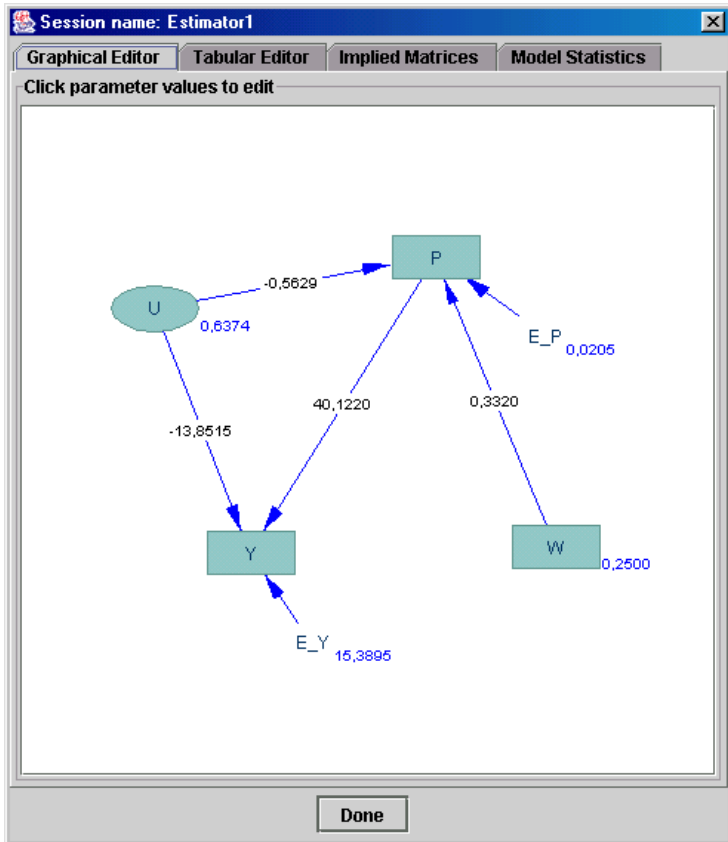
Graf na Rysunku 5.3 zgodny jest z założeniem Markowa stosowanym w podejściu BN. Oparte na założeniu wierności algorytmy przeszukiwania przestrzeni rozwiązań nie prowadzą do poprawnego rezultatu, gdyż brak jest odróżnienia rodzajów danych, zwłaszcza w zaproponowanej tu postaci wyróżnienia grup danych interwencyjnych względem różnych rodzajów oddziaływań.

Podobny problem program TETRAD napotyka w drugim przykładzie. Również tutaj algorytmy uchylające założenie o przyczynowej wystarczalności, FCI i GES, nie prowadzą i nie mogą prowadzić do odkrycia grafu, który Lauritzen(2004) zaproponował w komentarzu do wypowiedzi Rubina (2004).

Graf wskazany przez Lauritzena dla obu zaprezentowanych przez Rubina przykładów – z estymowanymi przez program TETRAD współczynnikami oraz wprowadzoną przez Lauritzena zmienną ukrytą – przedstawia Rysunek 5.4. Rysunek 5.5 przedstawia taki graf dla pierwszego z przykładów Rubina.

Brak krawędzi między U a W na obu rysunkach ma, zdaniem Lauritzena, reprezentować rodzaj przyjętego planu badawczego. W opisywanych przez Rubina badaniach nad szczepionką na wąglika dobór wolontariuszy do grupy poddanej oddziaływaniu oraz dobór makaków do grupy poddanej oddziaływaniu są losowe. Tym samym więc wszystkie nieuwzględnione

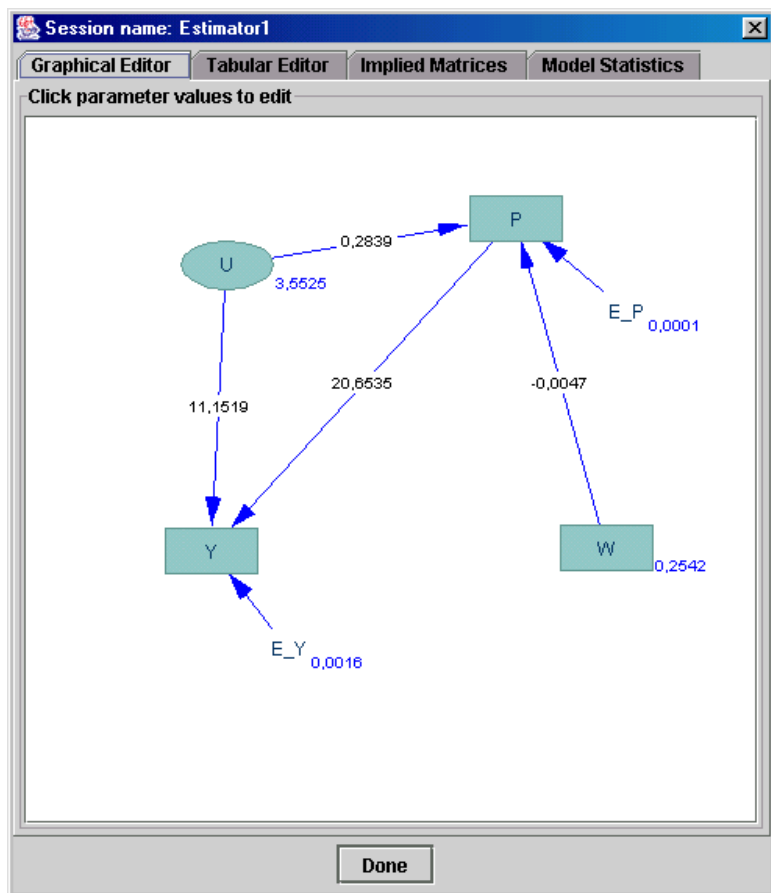
jawnie w modelu oddziaływania przyczynowe na członków grupy poddanej oddziaływaniu oraz niepoddanej oddziaływaniu są takie same. Z tego planu badawczego wynika więc odrzucenie w modelu jednego czynnika, który obciążałby zaobserwowane wartości zmiennej W .



Rysunek 5.4. Model graficzny z estymowanymi parametrami dla danych zaobserwowanych z obu przykładów dla szczepionki wąglika.

W odpowiedzi na komentarz Lauritzena Rubin zwraca uwagę na wieloznaczność reprezentacji graficznej. Brak krawędziw niektórych przypadkach, np. między U a W , ma oznaczać brak istnienia zależności przyczynowej między dwoma zjawiskami (poziom określany przez Rubina jako „the science”), ale w innych przypadkach oznacza sposób planowania badań tych

zależności lub przypisania do określonego oddziaływania. Ta wieloznaczność okazuje się zasadniczym problemem przy próbie uogólnienia lub ekstrapolacji otrzymanych zależności: które z przedstawionych graficznie zależności należy potraktować jako związane z *aktualnym* badaniem i sposobem *jego* przeprowadzenia, a które należy potraktować jako relacje między zjawiskami niezależne od szczegółów określonego badania.



Rysunek 5.5. Model graficzny z estymowanymi parametrami dla danych zaobserwowanych z pierwszego przykładu dla szczepionki węglika.

Podsumowując tę dyskusję, podkreślę, że podejście do graficznego modelowania zależności przyczynowych BN natrafia w analizie podanych

przez Rubina przykładów na dwie zasadnicze trudności: niemożliwość odkrycia modelu dla danych pochodzących z populacji zróżnicowanych z uwagi na nierejestrowane w badaniach zjawisko (zmienna ukryta) oraz wieloznaczność reprezentacji zależności przyczynowych, które można interpretować jako istniejące niezależnie od aktualnego badania lub wynikających z realizowanego planu badawczego.

Te dwie zasadnicze trudności nie dotyczą podejścia proceduralnego. Jak wskazywałem wcześniej, odkrycie niezgodności między wynikami stosowania kryterium proceduralnego do danych z różnych próbek/populacji prowadzi do jednego z dwóch wniosków. Może zaistnieć sytuacja, gdy analizowane dane zostaną uznane za dane interwencyjne, ale względem *różnych* oddziaływań. W związku z tym połączenie ich tworzy zbiór danych nieinterwencyjnych, w stosunku do którego można jedynie stwierdzić zachodzenie zależności przyczynowych względem modelu.

Druga sytuacja dotyczy przypadków uchylecia założenia o wystarczalności przyczynowej. Jeśli w takim przypadku stwierdza się niezgodności w analizie korelacji cząstkowych w zaobserwowanych danych, wówczas kryterium proceduralne w ogóle nie może być poprawnie zastosowane. Dla określenia zachodzenia zależności przyczynowych w danej dziedzinie wymagane jest jawne uwzględnienie zjawisk zmieniających zaobserwowane korelacje cząstkowe dla różnych fragmentów tej dziedziny (populacji, próbek). Wskazanie niezgodności w takiej analizie jasno pokazuje, że ten wymóg nie został spełniony. Rezygnując z jawnej reprezentacji takiego zjawiska lub zjawisk, można w dalszym ciągu przeprowadzać analizy zależności cząstkowych, posługując się kategorią zmiennej ukrytej. Wyniki takiej analizy są jednak uzależnione od założeń wkomponowanych w konstrukcję modelu. Dlatego też w tym przypadku można co najwyżej stosować kryterium proceduralne dla zależności przyczynowych względem modelu.

CZĘŚĆ II

ZARYS METODOLOGII PROCEDURALNEJ

6

Projekcyjność praw naukowych

Na proponowany w tej części zarys metodologii proceduralnej złożą się dwa kluczowe dla tej problematyki zagadnienia, a mianowicie koncepcja wyjaśniania¹ jako subsumpcji proceduralnej, której integralną częścią – a zarazem drugim zagadnieniem – jest koncepcja roli eksplanacyjnej projekcyjnych uogólnień modeli przyczynowych, spełniających kryterium proceduralne.

Rozważania rozpoczną od tematyki praw, by określić pojęcie i rolę eksplanacyjną uogólnień projekcyjnych, a następnie, aby poruszyć temat koncepcji wyjaśniania. Staram się przy tym przestrzegać następującej zależności i rozgraniczenia kompetencji między metodologią a filozofią nauk:

[...] rozważania filozoficzne, których przedmiotem jest poznanie naukowe, i rozważania metodologiczne przebiegają jak gdyby na dwóch odrębnych, choć równoległych płaszczyznach. Rozważania filozoficzne mają znacznie bardziej ogólny charakter. [...] Ta różnica pociąga za sobą kolejną i może w praktyce donioślejszą różnicę między filozofią i metodologią – techniki badawcze stosowane przez obie nauki są wyraźnie odmienne. Współczesna metodologia w coraz szerszym zakresie stosuje techniki formalne. [...]

A jednak mimo różnic, [...] metodologia i filozofia [...] są w istotny sposób od siebie zależne. Prześledzenie jakichkolwiek obszerniejszych rozważań metodologicznych ujawnia istnienie wielu założeń natury filozoficznej, na jakich rozważania te są oparte. Z drugiej strony koncepcje filozoficzne, ze względu na ich bardzo ogólny charakter, mogą być weryfikowane tylko wówczas, gdy znajdują zastosowanie w rozważaniach bardziej specjalistycznych, a więc np. w metodologii. (Wójcicki 1982, 15)

¹ W niektórych koncepcjach nauki wyjaśnianie jest traktowane jako jedyna kluczowa kategoria, np. J. Kmita i L. Nowak piszą: „Struktura nauki jest strukturą wyjaśniania” – dodajmy: w sensie omawianej w Punkcie 8.1 koncepcji DN; (1968, 91). Por. też (Kim 1999).

Istnieje wiele sposobów uogólnienia tego, co obserwujemy w doświadczeniu. Nieliczne określane są mianem praw naukowych.² Co o tym decyduje? Metodolodzy i filozofowie nauki proponują liczne kryteria, np. uniwersalność, bezwyjątkowa obowiązywalność, zawartość informacyjna, prostota czy płodność w generowaniu wyjaśnień.³

Przedmiotem rozważań w tym rozdziale jest projekcyjność, jako jedna z cech charakteryzujących prawa naukowe.⁴ Nie stanowi ona podstawy do kompletnej charakterystyki praw naukowych. Moje zadanie polega na wykazaniu, że projekcyjność jest cechą wystarczającą dla zdeterminowania roli eksplanacyjnej uogólnień, a w szczególności praw naukowych.

Konsekwencją tego stanowiska jest zasadnicza rewizja tradycyjnej koncepcji wyjaśniania naukowego, w której centralną rolę eksplanacyjną przypisuje się właśnie prawom. Ta rewizja jest szczególnie istotna dla statusu metodologicznego nauk, w których trudno byłoby podać uogólnienia spełniające kryteria tradycyjnie podawane dla praw naukowych. Do tych tematów powrócimy w trakcie dyskusji zagadnienia wyjaśniania w dalszych rozdziałach tej części.

Zasadnicze zadanie metodologii proceduralnej upatrywałbym w wyodrębnieniu klasy *uogólnień projekcyjnych*. Wskazanie zaś, które z nich spełniają ponadto kryteria prawa naukowego, jest tu zadaniem mniej istotnym i, jak sądzę, niezależnym od ich funkcji eksplanacyjnej. Tym samym proponowana przeze mnie koncepcja metodologii proceduralnej wyznacza pośrednie stanowisko między koncepcją dominacji praw (jako jedynej dopuszczalnej podstawy wyjaśnień naukowych) a koncepcją sceptyczną (według której prawa naukowe są zbędne w nauce).

² Ponieważ nie podejmuję w tej pracy kwestii ontologicznych dotyczących praw, więc „prawa nauki” traktuję niekiedy zamiennie z „prawami przyrody”, co nie pogwałca kanonu „Nie musimy jednak wciąż o tym rozróżnieniu pamiętać, zajmując się zagadnieniami metodologii nauki” (Krajewski 1998, 18). Por. też (Carroll 1987). R. Wójcicki odróżnia prawidłowości od praw – termin „prawo” odnosi wyłącznie do praw naukowych; por. (1982, 71).

³ Bogatą listę kryteriów praw naukowych – i w niektórych punktach odmienną od tu przyjętej – podał S. Mazierski (1995, 97-123). Wydaje się jednak, że wśród nich eksponował właśnie ich projekcyjność (1995, 83). Podobne stanowisko przyjmował także Amsterdamski, proponując zbliżone do projekcyjności pojęcie mocy prognostycznej (1983, 105-107).

⁴ Zwięzły zarys dziejów pojęcia prawa oraz wprowadzenie do problematyki w polskiej literaturze podają S. Mazierski (1995, 13-21) i W. Krajewski (1998, 12-13).

6.1 Koncepcja projekcyjności uogólnień

Dyskusję nad kwestią projekcyjności uogólnień zainicjował N. Goodman w swojej krytyce formalnej teorii uogólnień indukcyjnych, jaką w latach 1940. i 1950. skonstruował R. Carnap.⁵ Goodman starał się wykazać, że opis danych doświadczenia tak samo potwierdza dwa niezgodne ze sobą uogólnienia tych danych, a w szczególności takie, które prowadzą do sprzecznych przewidywań.

Ilustruje to specjalnie skonstruowany przez Goodmana predykat „ziebieski”: x jest ziebieski wtw x jest zielony do czasu t i jest niebieski później. O ile t oznacza punkt czasu późniejszy niż teraźniejszość, to wszystkie zaobserwowane dotąd szmaragdy można opisać zarówno jako zielone, jak i – wykorzystując predykat Goodmana – jako ziebieskie. Jeśli jednak zapytać o to, jaki kolor będzie miał szmaragd znaleziony później niż t , oba uogólnienia prowadzą do niezgodnych przewidywań.

Tylko jedno z tych dwóch przewidywań można określić, zdaniem Goodmana, jako projekcyjne. „Wszystkie szmaragdy są zielone” jest semantycznie symetrycznym uogólnieniem względem „Wszystkie szmaragdy są ziebieskie”. Można je bowiem wzajemnie definiować, wprowadzając jako pierwotny np. predykat ziebieski i za jego pomocą definiując predykat zielony: x jest *zielony* wtw x jest ziebieski do czasu t , a później jest nielony.

Goodman odrzucał możliwość eksplikacji asymetrii między predykatami zielony a ziebieski na gruncie ontologicznym. Twierdził natomiast, że jest to różnica wyłącznie z poziomu pragmatyki semiotycznej. Wynika ona z faktu, że w dotychczasowych uogólnieniach doświadczenia ludzie posługiwali się właśnie predykatem zielony, a nie ziebieski. Stąd ten pierwszy, a nie drugi, należy uznać za *zakorzeniony* w praktyce językowej.

W charakterystyczny dla siebie sposób Goodman wiąże więc projekcyjność z własnościami językowymi predykatów występujących w uogólnieniach. W sformułowaniu „paradoksu ziebieskich” Goodman implicite zakłada jednak wiedzę ogólną. Rzekoma paradoksalność jest bezpośrednią konsekwencją wieloznaczności tego, co w różnych przypadkach traktowane jest jako składnik przyjmowanej wiedzy ogólnej.⁶

Jeśli przyjąć, że ta wiedza ogólna zawiera informacje pozwalające na wprowadzenie asymetrii między predykatem zielony a ziebieski, wówczas

⁵ Wprowadzenie historyczne do tej dyskusji podaje Kawalec (2003a, 129-33). Por. też (Scheffler 1967; 1982).

⁶ Szczegółowo zagadnienie to dyskutuje Kawalec (2001).

nie ma podstaw do sformułowania paradoksu. Na tej samej podstawie można bowiem pokazać, że predykat ziebieski jest nowotworem językowym, zdefiniowanym w podstawowych kategoriach zielony i niebieski.

Jeśli z kolei przyjąć, że wiedza ogólna nie jest wystarczająca do wprowadzenia asymetrii między tymi dwoma predykatami, wówczas nie ma podstaw do wystąpienia paradoksu. Powstaje oczywiście problem wybrania jednego z co najmniej dwóch możliwych uogólnień, jednak nie jest to już paradoks. Można to wyrazić, zastępując predykaty zielony i ziebieski wyrażeniami, które nie presuponują asymetrii w wiedzy ogólnej, np. każdy szmaragd należy do klasy A albo każdy szmaragd należy do klasy B. Wybór jednego z tych dwóch uogólnień – bez założenia jakiegokolwiek wiedzy ogólnej – nie ma w sobie nic paradoksalnego. Przy braku założeń wydaje się to niewykonalnym zadaniem.

Paradoksalność powstaje jako złączenie tych dwóch przypadków w jeden: wydaje się, że oba uogólnienia – mimo że dają niezgodne przewidywania – są tak samo potwierdzone przez doświadczenie, a jednocześnie jest rzeczą wiadomą, które z nich należy preferować. W tym przypadku jednak w grę wchodzi *dwa różne* rozumienia wiedzy ogólnej.

Wybranie jednego z możliwych uogólnień jest trudnym problemem. Jako problem, a nie paradoks, może być jednak – i jest faktycznie w nauce – rozwiązywany. Najbliższą jego postacią, często spotykaną w literaturze naukowej, jest problem dopasowania formy zależności funkcyjnej między zmiennymi na podstawie dostępnych danych.

Dla rozwiązywania tej trudności w konkretnych przypadkach opracowano wiele narzędzi, m.in. bayesowską teorię wnioskowania sekwencyjnego (por. Grabski i Jaźwiński 2001) czy kryterium Aikaike (Burnham i Anderson 2002). Niezależnie od wybranego rozwiązania, przyjęcie pewnych założeń w punkcie wyjścia wydaje się nieodzowne.⁷

Powracając do propozycji Goodmana, zwróćmy uwagę, że sformułował on teorię projekcyjności w kategoriach zakorzenienia w praktyce językowej, wskazując jednoznacznie i bez wątpliwości predykat zielony jako predykat projekcyjny. Tym samym dał dowód temu – co podkreśliłem wyżej – że faktyczne rozstrzygnięcie projekcyjności predykatu dokonuje się w przyjętej wiedzy ogólnej.

Zastosowane implicite przez Goodmana rozwiązanie, z istotną wszakże modyfikacją, przyjmę w proponowanej tu koncepcji projekcyjności. Mody-

⁷ Rozwiniętą argumentację potwierdzającą ten wniosek przedstawił E. Sober (1994).

fikacja, o której mowa, wynika przede wszystkim z faktu, że podany przeze mnie zarys koncepcji projekcyjności jest ograniczony do uogólnień stosowanych w nauce, a przede wszystkim w badaniach nieeksperymentalnych. Stąd też jako zasadniczy element wiedzy, którą można założyć w momencie formułowania projekcji, wskażę model przyczynowy, spełniający wprowadzone w Punkcie 2.3 kryterium proceduralne.

Precyzyjniejsze sformułowanie projekcyjności poprzedzę kilkoma istotnymi pytaniami. Goodman'a teoria zakorzenienia językowego wyznacza jeden z dwóch zasadniczych kierunków poszukiwania charakterystyki projekcyjności. Alternatywny kierunek wyznaczył esej W. Quine'a „Natural kinds” (1970). Pomijając przyjęty przez Quine'a redukcjonizm rodzajów naturalnych, zwróćmy uwagę, że zasadniczy element tej koncepcji to poszukiwanie ontologicznego ugruntowania dla asymetrii między predykatami projekcyjnymi a nieprojekcyjnymi. O szmaragdach mówimy więc w kategoriach zielony, a nie ziebieski, dlatego, że tylko pierwszy z tych predykatów odpowiada istniejącemu niezależnie rodzajowi naturalnemu.

Wyeksplikowane w części pierwszej przesłanki przyjętego przeze mnie podejścia proceduralnego skłaniają do rezerwy wobec tego rozwiązania. Wymaga ono przyjęcia w punkcie wyjścia bardzo mocnych założeń, które – jak pokażę niżej – są zbędne przy charakterystyce pojęcia projekcyjności.

Ponadto każda próba uniknięcia arbitralności takich założeń nieuchronnie wydaje się prowadzić do naturalizmu,⁸ chociaż może niekoniecznie do jego redukcyjnej postaci przyjętej przez Quine'a (1970).⁹ Jeśli bowiem uzależnić dokonywane w nauce projekcje od tego, że np. *X* jest rodzajem naturalnym, wówczas eksplikacja tego warunku w formie redukcyjnej lub niereduksyjnej doprowadzi do reprezentacji tego faktu w danej teorii naukowej. Stąd zaś blisko do twierdzenia, że nauka reprezentuje nie tylko rodzaje naturalne, ale także inne kategorie ontologiczne. Ostatecznie zaś, że to nauka miałaby przesądzać o tym, jakie jest ontologiczne umebłowanie świata.

⁸ Wprowadzone przez S. Judyckiego (1995) odróżnienie uwarunkowania od związku przyczynowego systematycznie w obronie antynaturalizmu w filozofii umysłu rozwija Żegleń (2003, 185-201).

⁹ Taką niereduksyjną wersję wydaje się przyjmować L. Gumański: „pojęcie związku przyczynowego jest podrzędne wobec pojęcia związku naturalnego (zachodzącego w rzeczywistości). [...] Kiedy mówimy o oddziaływaniu przyczyn, o wywoływaniu skutków, mamy na myśli zależność zachodzącą w naturze, wykluczamy zaś związki umowne, konwencje” (2004, 114).

W tej pracy nie podejmuję dyskusji z naturalizmem.¹⁰ Szkic argumentacji naturalistycznej, który podałem, ma służyć jedynie unaocznieniu faktu, że przyjęcie zaproponowanej przez Quine'a strategii opracowania koncepcji projekcyjności uwikłałoby mnie w bardzo złożone dyskusje filozoficzne.¹¹ A tym samym uzależniałoby koncepcję projekcyjności od przyjętego rozstrzygnięcia w jednej z najbardziej dziś kontrowersyjnych filozoficznie kwestii.

Wybieram więc możliwość podania koncepcji projekcyjności na potrzeby metodologii nauk z pominięciem tych kontrowersji. Nie oznacza to, że filozoficzna kwestia ustalenia ontologicznych podstaw projekcyjności nie może być studiowana równolegle. Obfitość problematyki podejmowanej w tej pracy skłania jednak do pozostawienia tej kwestii do osobnego studium.

Przy konstruowaniu koncepcji projekcyjności nie można także ominąć problemu – na który zwrócił uwagę T. Kuhn – rozbieżności schematów pojęciowych czy paradygmatów, jakie mogą różnić naukowców. Jeśli bowiem projekcja modelu przyczynowego byłaby względna w stosunku do jednego schematu pojęciowego i niedostępna dla posługujących się innym schematem pojęciowym, wówczas projekcyjności należałoby przypisać nikłą rolę w uprawianiu nauki i szerzej – w metodologii nauk.

Rozbieżność schematów pojęciowych, na którą zwracał uwagę Kuhn, dobrze jest to zilustrowane przez detalicznie omówiony w części pierwszej przykład badań nad cholerą w XIX wieku. Przywołajmy kilka kluczowych rozbieżności. Ustanowiona z ramienia parlamentu brytyjskiego the Board of Health miała za zadanie zidentyfikowanie przyczyny epidemii w Londynie, aby ustanowić środki zapobiegawcze. Działający niezależnie lekarz Snow realizował podobne zadanie.

Komisja przyjmowała miazmatyczną koncepcję cholery, scharakteryzowaną w Punkcie 2.2 jako paradygmat losowy: do wybuchu epidemii dochodzi wskutek natrafienia oparów z Tamizy zawierających truciznę cholery na sprzyjające warunki, np. nieczystości pochodzenia organicznego, zanieczyszczona woda, nieuregulowane ścieki. Wówczas dochodzi do „skatali-

¹⁰ Ze współczesnych polemik z antynaturalizmem zwraca uwagę U. Żegleń systematyczna argumentacja w filozofii umysłu, oparta na filozoficznej analizie pojęcia przyczynowości (Żegleń 2003, 185-201).

¹¹ Ponadto należałoby też zająć systematyczne stanowisko w szczegółowej, ale żywo dziś dyskutowanej, kwestii przyczynowości mentalnej, a zwłaszcza zainspirowanej anomalizmem Davidsona dyskusji wokół istnienia ścisłych praw w psychologii. Zwięźle prezentuje je Żegleń (2003, 177-84).

zowania” tej trucizny i zamienienia jej w „czynnik aktywny”, czego przejawem jest niespodziewany wybuch epidemii.

Dane komisji zbierane były w celu potwierdzenia tej hipotezy. Stąd też znaczącą część raportu zajmują dane meteorologiczne i pomiary bezpośrednio związane z Tamizą (np. jej temperatura, wilgotność powietrza itp.). Kontrolę poddawano również jakości powietrza i wody. Sprawdzano także jakość instalacji kanalizacyjnych.

Snow w swoich badaniach przyjmował deterministyczny punkt widzenia. Obserwacja objawów chorobowych u poszczególnych pacjentów doprowadziła go do przekonania, że przyczyną – „czynnikiem aktywnym” – jest niewidzialny pasożytniczy mikroorganizm, który wraz z wodą lub pożywieniem dostaje się do układu pokarmowego, tam po okresie inkubacji dokonuje zmian w składzie chemicznym krwi, co w większości przypadków prowadzi do śmierci. Dane, jakie Snow zebrał w celu potwierdzenia tej hipotezy, miały wykazać przede wszystkim, że woda jest „czynnikiem aktywnym”, a nie – jak przyjmowała to komisja – „czynnikiem sprzyjającym”.

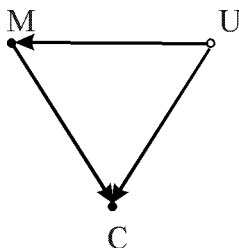
Te dwa radykalnie odmienne sposoby patrzenia na zjawisko cholery nie zniknęły wraz z *On the mode of communication of cholera* Snowa, pracą demonstrującą wpływ skażonej wody na zachorowalność. Podczas piątej fali epidemii cholery w 1884 roku (kiedy to Koch wyizolował *vibrio cholerae*) nadal koncepcja miazmatyczna była szeroko rozpowszechniona (Sack i in. 2004, 223). Co więcej, w czasie badań prowadzonych równoległe do Snowa, Hassall, jeden z członków komisji, zaobserwował *vibrio cholerae*, jednak, kierując się przyjętą perspektywą, zinterpretował obecność tych bakterii jako kolejny przejaw działania trucizny cholery. Ten przykład dobitnie ilustruje „niewspółmierność” tych dwóch perspektyw, nawet na tak podstawowym poziomie, jakim jest poziom obserwacyjny.

Z drugiej zaś strony warto podkreślić, że zalecenia podawane na gruncie perspektywy losowej okazywały się bardzo skuteczne w zapobieganiu epidemii cholery. Przede wszystkim wprowadzono regulacje prawne w kwestii instalacji kanalizacyjnych, a później także w sprawie filtrowania wody. Ponadto współczesne badania potwierdzają jedną z tez komisji, a mianowicie, że okresowe nasilenie się epidemii cholery jest „związane z wodą”. Dokładniej – okresowy rozwój kolonii bakterii *vibrio cholerae* związany jest z fazami rozwoju planktonu oceanicznego i innych organizmów, które są nośnikami tych bakterii w wodzie oceanicznej i morskiej (Colwell 1996, 2030-31).

Przykład rozbieżności między przyjętą przez komisję perspektywą badawczą oraz a perspektywą przyjętą przez Snowa pokazuje, że projekcyjność nie jest jednak zrelatywizowana do określonej perspektywy. Co więcej, pozwala na taką ewolucję odmiennych perspektyw, która ostatecznie wyłoniła znaną nam współcześnie, a łączącą elementy tych – wydawałoby się – zupełnie niewspółmiernych schematów pojęciowych.¹²

Zasadniczą różnicę między perspektywą Snowa a perspektywą komisji widziałbym w tym, do jakiego rodzaju projekcji prowadziła każda z nich. Snow na podstawie badań poszczególnych pacjentów, przyjmując założenie o jednorodności biologicznej, dokonał projekcji prospektywnej: U każdego z dwóch ludzi A i B, których różni tylko to, że do przewodu pokarmowego A został wprowadzony pasożytniczy mikroorganizm (M), zaobserwujemy różnicę w tym, że A ujawni objawy choroby (C) po okresie inkubacji tego mikroorganizmu (tj. w przeciągu 1-5 dni). Taka projekcja prowadzi do sformułowania modelu przyczynowego ze zmienną ukrytą.

Próbka, którą dysponował Snow, nie pozwalała mu na stwierdzenie, czy nie istnieje w tym przypadku czynnik osłabiający wniosek. Nie była to próbka losowa, więc obserwowane objawy chorobowe mogły wynikać np. z poziomu odporności organizmu lub czynnika zaburzającego skład chemiczny krwi jako konfundera zależności między obserwowanym wprowadzeniem skażonego pokarmu a objawami choroby.



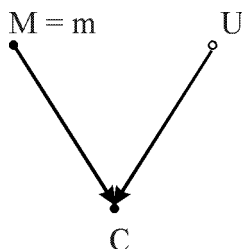
Rysunek 6.1.

Kolejny więc krok w postępowaniu Snowa polegał na podjęciu takich interakcji w oparciu o ten model, które pozwoliłyby na zastosowanie kryterium proceduralnego. Takim przykładem są liczne podawane przez niego przypadki obserwacji, kiedy wpływ zmiennej ukrytej $U \rightarrow M$ jest poddany

¹² Pomijam w tej pracy kwestie socjologiczne i psychologiczne, dotyczące mechanizmu powstawania rozbieżnych perspektyw oraz ich utrzymywania, nawet przy niezgodnych danych doświadczenia.

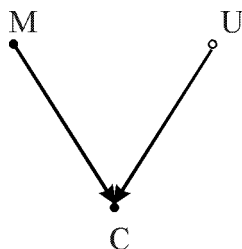
kontroli, a więc np. dwa osiedla niemal jednorodne pod względem położenia i warunków atmosferycznych (oraz innych nierejestrowanych w modelu oddziaływań) różniło tylko źródło wody.

Przykładem interakcji, który wzbudził duże zainteresowanie, było zdjęcie ramienia z pompy na Broad Street, co można potraktować jako interakcję zgodnie ze zmodyfikowaną postacią modelu z Rysunku 6.2.



Rysunek 6.2.

Zasadniczy, jak się podkreśla, krok w rozumowaniu Snowa polegał na dokonaniu kolejnej projekcji: U każdych dwóch grup ludzi A i B, których różni tylko to, że A korzysta ze skażonego źródła wody, zaobserwujemy nielosową różnicę w tym, że A ujawni więcej przypadków zachorowań na cholere.



Rysunek 6.3.

W tym przypadku brak krawędzi między M a U reprezentuje fakt, że próbka jest dobrana losowo. Oznacza to, że prawdopodobieństwo poddania oddziaływania M jest takie samo dla każdej z osób poddanych różnym innym nieuwzględnionym w tym modelu oddziaływaniami. Zmienna U , którą może być np. nieobserwowany przez Snowa poziom odporności lub czynnika modyfikującego chemiczny skład krwi, zmienia oddziaływanie ska-

zonej wody.¹³ Ustalenia Snowa pozwoliły na wykorzystanie eksperymentu naturalnego, opisanego bliżej w Punkcie 1.6, do stwierdzenia niezmienniczości zależności przyczynowej między spożyciem skażonej wody a zachorowalnością na cholere.¹⁴

Projekcje dokonywane przez komisję były zasadniczo retrospektywne. Hipoteza o losowym krążeniu trucizny cholery w atmosferze pozwalała w większości wypadków na wyjaśnienie, dlaczego została skatalizowana w ściśle określonym miejscu. Często dotyczyło to południowych dzielnic Londynu, gdzie spożywanie skażonej wody było skorelowane z ogólnie gorszymi warunkami życiowymi i większym zanieczyszczeniem. Znane były także przypadki, jak np. wdowa zamawiająca wodę z pompy przy Broad Street, lecz mieszkająca daleko, które komisja mogłaby wyjaśnić jedynie działaniem zmiennych ukrytych lub konfunderów nieuwzględnionych w tym modelu.

Projekcje prospektywne, gdyby takie sformułować z perspektywy przyjętej przez komisję, byłyby zasadniczo zgodne z całą klasą modeli przyczynowych zgodnych z zaobserwowanymi danymi. Przyjęcie hipotezy o losowym przenoszeniu się nieskatalizowanej trucizny cholery w zasadzie uniemożliwiało pozyskanie jakichkolwiek danych interwencyjnych dla uzyskania modelu przyczynowego. Brak takich danych uniemożliwiał z kolei nie tylko zastosowanie kryterium proceduralnego, ale ostatecznie podanie wyjaśnienia przyczyn epidemii cholery.

Wracając zatem do zaproponowanej przez Goodmana strategii rozwiązywania zagadnienia projekcyjności, zwrócę uwagę w analizowanym przykładzie na rzecz następującą. O rodzaju projekcji decyduje przyjęty model przyczynowy. Nie przesądza on o tym, czy sformułowana w oparciu o niego projekcja okaże się prawdziwa. Jest jednak decydujący pod względem umożliwienia zastosowania kryterium proceduralnego, które może prowadzić do zweryfikowania projekcji. Kolejnym więc elementem, na jaki należy zwrócić uwagę przy konstrukcji pojęcia projekcyjności, jest to, czy umożliwia ono sformułowanie takiego modelu, dla którego można zebrać dane interwencyjne i zastosować kryterium proceduralne. Podkreśłmy, że ta cecha

¹³ W ten sposób można wytłumaczyć różnicę w przebiegu choroby u różnych pacjentów.

¹⁴ W tym przypadku U może reprezentować niewyjaśnione na gruncie rozważanych przez Snowa zjawisk okresowe (sezonowe) nasilenie się zachorowań na cholere.

projekcji jest niezależna od przebiegu interakcji z badaną rzeczywistością i tego, czy jej wyniki okażą się zgodne z przewidywaniami, czy też nie.¹⁵

Podsumowaniem tych rozważań jest następująca propozycja rozumienia projekcyjności.

DEFINICJA 6.1. $F : \rightarrow G$ jest projekcją zależności przyczynowej $X \rightarrow Y$ względem modelu $M(\langle V, L \rangle)$ wtw $D \setminus D' \neq \emptyset$ oraz model M' , powstały z M przez zastąpienie pokrycia Markowa $bl(\{X, Y\})$ przez $bl(\{F, G\})$ względem zmiennych w V , spełnia kryterium proceduralne w D (istnieje działanie A , względem którego istnieją dane interwencyjne dla $F \rightarrow G$);

gdzie D' to dziedzina dla F, G , D dziedzina dla $X, Y \in V$.

Zależność $F : \rightarrow G$ w modelu M' może być uznana za przyczynową, wtedy gdy dla danych interwencyjnych minimalnego lub ekspandowanego pola interwencji $F \rightarrow G$ spełnione jest kryterium proceduralne.

Spróbuję teraz odnieść tę definicję do przykładu podanego przez Goodmana. W tym przypadku, upraszczając, można przyjąć, że dla pewnej skończonej grupy zaobserwowanych szmaragdów w D wiedza ogólna zawiera zależność przyczynową $S \rightarrow Z$ w tym sensie, że struktura tego kamienia szlachetnego jest przyczyną fenomenologicznie rejestrowanej przez ludzi własności.

Za projekcję Goodman uznaje uogólnienie o postaci: $S : \rightarrow Z$ dla D' zawierającego klasę wszystkich kamieni o strukturze S . Natomiast uogólnienie $S : \rightarrow G$ (gdzie G oznacza ziebieski) nie jest w tym przypadku projekcją. Na podstawie Definicji 6.1 tę różnicę można wyjaśniać następująco. Dla zależności $S : \rightarrow Z$ istnieje, zgodnie z modelem M zawierającym zależność $S \rightarrow Z$, czynność A (w polu interwencji $S \rightarrow Z$ lub jego ekspansji) taka, że posiadamy dane interwencyjne. Nawet jeśli trudno byłoby podać precyzyjny opis takiej czynności (np. wnioskowania na podstawie serii eksperymentów naturalnych), to w przeciwieństwie do drugiego z uogólnień taki opis jest niesprzeczny logicznie.

Istnienie takiej czynności jest natomiast wykluczone dla $S : \rightarrow G$ na gruncie modelu M . Warunkiem tego, by model M' , do którego należy zależność $S \rightarrow G$, był w tym samym wzorcu co M , taka czynność musiałaby zawierać także kontrolę zmiennej wchodzącej w skład G , a mianowicie czasu. Nie można jednak podać niesprzeczego opisu takiej czynności, jaka wymaga porównania efektów zmian dokonywanych na S w różnych okresach czasu,

¹⁵ Z innej niż przyjęta tu perspektywy o relacji między proceduralnością badań a projekcyjnością ich wyników pisał M. Przełęcki (1959/1966, 120).

ale niezależnie od czasu. Zgodnie więc z przyjętą definicją projekcyjności uogólnienie $S \rightarrow G$ nie spełnia tego warunku.

Jeśli dynamikę postępowania badawczego charakteryzować w sposób proceduralny, to otrzymamy schemat kroków badawczych:

interakcja z rzeczywistością \rightarrow model przyczynowy \rightarrow projekcja \rightarrow
interakcja ...

W Y J A Ś N I A N I E

Rysunek 6.4.

W tym schemacie, zgodnie ze strategią przyjętą w części pierwszej, zasadniczą rolę odgrywa – wprowadzona w tym studium jako pojęcie pierwotne – interakcja badawcza z rzeczywistością.

Ta pierwotność ujawnia się także w koncepcji sądów kontradiktorycznych. Standardowe dziś rozwiązanie polega na przyjęciu semantyki możliwych światów dla sądów kontradiktorycznych, a następnie wykorzystanie ich w eksplikacji tak doniosłych dla metodologii pojęć, jak przyczynowość czy prawo naukowe.¹⁶

Zgodnie z proponowanym tu stanowiskiem proceduralnym takie rozwiązanie nie może jednak zostać przyjęte. Podałem bowiem eksplikację kategorii przyczynowości odwołując się zasadniczo do pojęcia interakcji z badaną rzeczywistością jako pierwotnego oraz do danych interwencyjnych. Warunki prawdziwości zdań kontradiktorycznych można więc jedynie wprowadzić, podając odniesienie do modelu przyczynowego poprzez *introjekcję*, czyli operację odwrotną do zdefiniowanej wyżej projekcji.¹⁷

Zilustruję to przykładem. Zdanie kontradiktoryczne: „Gdyby profesor ornitologii wypił w dniu powrotu do domu wodę z pompy przy Broad Street, nie zważając na jej nieprzyjemny zapach, to zachorowałby na cholere”. To zdanie kontradiktoryczne można uznać za prawdziwe z uwagi na model przyczynowy, jaki przedstawia Rysunek 6.1. Przy dokonywaniu takiej introjekcji

¹⁶ Por. (Lewis 1973; 1986a; 1986b; Collins i in. 2002)

¹⁷ Nie proponuję tu teorii zdań kontradiktorycznych, a jedynie sugestią tego, jak można by taką teorię sformułować w sposób zgodny z koncepcją proceduralną.

konieczne jest oczywiście przyjęcie określonych założeń, które szczegółowo eksplikuje np. Rubin (1986).

Pozostałą część tego rozdziału poświęcę dyskusji nad tym, jaka zależność zachodzi między projekcyjnością uogólnień a cechami przypisywanymi prawom naukowym. Zasadnicze pytanie dotyczy tego, czy takie tradycyjnie przypisywane prawom cechy, jak uniwersalność, informatywność i prostota czy unifikacyjność, są warunkami koniecznymi pełnienia przez uogólnienia projekcyjne roli eksplanacyjnej.¹⁸ Przedłużeniem tej argumentacji jest wykazanie w Punkcie 6.2, że przyjęcie projekcyjności jako jedyne go warunku koniecznego dla roli eksplanacyjnej uogólnień pozwala uniknąć szeregu trudności, na jakie natrafiają koncepcje praw naukowych.¹⁹

6.2 Projekcyjność uogólnień a uniwersalność praw naukowych

Zastanawiając się nad związkiem między właściwościami praw naukowych a ich rolą eksplanacyjną, przyjmę uproszczoną koncepcję tej drugiej. Na obecne potrzeby taka jej wersja będzie wystarczająca, natomiast jej szczegółową postać omawiam w Rozdziale 8 i 9.

Na czym więc polega wyjaśnienie jakiegoś zdarzenia lub regularności zjawisk? Istotnym elementem wyjaśnienia jest podanie ogólnej zależności, której szczególną realizacją w danej sytuacji jest to zdarzenie lub regularność. Sukces teorii Newtona polegał na tym, że prawo grawitacji, stanowiące jej zasadniczą część, było podstawą wyjaśnienia wielu zdarzeń i regularności obserwowanych na Ziemi i na niebie.

W skrócie, poprzez projekcję można uogólnić sformułowany dla danych interwencyjnych model przyczynowy na takie zdarzenie lub regularność, które nie zostały ujęte w danych, a do których stosowane było kryterium proceduralne. Posiadając z kolei dane interwencyjne dla projekcji modelu, można określić obowiązywalność w nim kryterium proceduralnego. Powstaje więc tym samym model przyczynowy, którego szczególnym przy-

¹⁸ Nie zakładam tu utożsamienia własności bycia prawem z funkcją eksplanacyjną, por. (Braithwaite 1955, 303).

¹⁹ Dla uniknięcia kontrowersji wokół istnienia praw w naukach społecznych, Amsterdamski stosuje podobną do zaproponowanej tu strategię, uzależniając rolę eksplanacyjną uogólnień w naukach społecznych od ich mocy prognostycznej. Tę ostatnią Amsterdamski charakteryzował w sposób zbliżony do omawianej tu własności projekcyjności uogólnień, por. (1983, 123-128).

padkiem jest zdarzenie lub regularność, dla których poszukujemy wyjaśnienia.

Czy takie tradycyjnie przypisywane prawom naukowym cechy, jak uniwersalność, prostota i informacyjność czy unifikacyjność, są nieodzownym elementem w wyżej scharakteryzowanym pobieżnie procesie wyjaśniania? Poniżej postaram się wykazać, że tak nie jest.²⁰

Rozpocznę od zagadnienia uniwersalności praw. J. Earman poddał tę cechę praw szczegółowej analizie (1978).²¹ Pomijając na razie kwestię przejścia od regionów czasoprzestrzeni do takich zjawisk, jak cholera czy palenie papierosów, przywołam wyniki analizy Earmana. Znane z literatury przedmiotu²² charakterystyki uniwersalności praw Earman grupuje następująco:

- 1) prawa nie odnoszą się do określonych lokalizacji w czasoprzestrzeni;
- 2) prawa mają nieograniczony zakres w czasie i przestrzeni;
- 3) prawa nie zawierają jawnie wymienionych współrzędnych czasowych lub przestrzennych;
- 4) prawa są niezmiennicze względem przekładów czasowo-przestrzennych.

Przy założeniu absolutnej (jednej i zafiksowanej) czasoprzestrzeni wśród tych czterech możliwych znaczeń, Earman wyróżnia 2), dla którego podaje słabszą i mocniejszą interpretację. W mocniejszej interpretacji uniwersalność prawa oznacza, że nie ma takiego niepustego fragmentu czasoprzestrzeni, w którym prawo byłoby spełnione w każdej możliwej sytuacji dla tego fragmentu. Oznaczałoby to bowiem zjawisko znane jako przepelnienie, czyli zgodność prawa z każdym możliwym stanem rzeczy w tym fragmencie. Mówiąc inaczej, prawo, nie wykluczając w tym fragmencie żadnej z możliwości, niczego nie regulowałoby. Taki więc fragment, gdyby istniał, byłby poza obowiązywalnością tego prawa. Dlatego też takie prawo nie byłoby uniwersalne.

Słabsza interpretacja dotyczy przypadku, gdy prawo formułuje zależność dotyczącą określonego fragmentu czasoprzestrzeni. Odniesienie do tego fragmentu może być – jak podkreśla Earman – sformułowane na jeden z

²⁰ Por. też (Czeżowski 1968, 201).

²¹ Analiza J. Earmana minimalizuje dodatkowe założenia, zwłaszcza ontologiczne, które mogą być podstawą trudności niespecyficznych dla pojęcia uniwersalności, zwłaszcza gdy w grę wchodziłoby pojęcie konieczności; por. (Amsterdamski 1983, 54-73; Such 1972, 143-63).

²² Por. m.in. (Braithwaite 1955; Goodman 1983; Hempel 1965; Reichenbach 1954; Schlick 1932/1949).

dwóch sposobów. Bądź zostanie wymieniona pewna stała, bądź zostanie podana charakterystyka tego fragmentu przez odwołanie się do zakrzywienia (całej) czasoprzestrzeni. W pierwszym przypadku możliwe jest więc scharakteryzowanie tego fragmentu lokalnie („pozycjonalne”), np. przez ostensję. Przy pozycjonalnym sformułowaniu na podstawie samego prawa można pokazać, że istnieje fragment czasoprzestrzeni, który go nie spełnia.

Nie jest tak natomiast w przypadku, gdy charakterystykę fragmentu czasoprzestrzeni podaje się poprzez odniesienie do zakrzywienia czasoprzestrzeni (nielokalnie). Wówczas na podstawie samego prawa nie można pokazać, że nie jest ono uniwersalne. To właśnie przesądza o słabszej interpretacji uniwersalności: uniwersalność prawa oznacza, że nie ma takiego niepustego fragmentu czasoprzestrzeni, o którym na podstawie samego prawa można byłoby dowieść, że byłoby ono w nim spełnione w każdej możliwej sytuacji.

Niezależnie od tego, którą z tych interpretacji wybierzemy – rozumienie uniwersalności w sensie wyróżnionym jako 1) prawa nie odnoszą się do określonych lokalizacji w czasoprzestrzeni – można podać przykłady sytuacji, gdy te interpretacje są spełnione, a nie jest 1). Wyobraźmy sobie prawo, które podaje czas rozpadu jądra atomowego dla dowolnego fragmentu czasoprzestrzeni jako jego relację do szczególnego regionu czasoprzestrzeni, np. wielkiego wybuchu. Takie prawo stosuje się uniwersalnie w sensie 2) (przy obu interpretacjach), ale nie w sensie 1). Dlatego też Earman odrzuca 1) jako eksplikację uniwersalności prawa naukowego.

Znaczenie podane jako 3) prawa nie zawierają jawnie wymienionych współrzędnych czasowych lub przestrzennych – ma w zamierzeniu podawać warunek dla 4), a więc dla niezmienniczości względem przekładów czasoprzestrzeni. Earman (1978, 176) pokazuje, że nie jest to jednak ani warunek konieczny, ani wystarczający. Dlatego też przechodzę do analizy 4). Po podaniu formalnej postaci tego rozumienia uniwersalności, Earman dowodzi, że uniwersalność w sensie 2) nie implikuje uniwersalności w sensie 4), ale w odwrotną stronę ta implikacja zachodzi dla niektórych ważnych przypadków (Earman 1978, 177). Stąd wniosek, że – przynajmniej w niektórych ważnych przypadkach – uniwersalność w sensie 4) jest wtórna wobec 2).

Rozważania Earmana prowadzą więc do wyróżnienia 2) jako sensu uniwersalności praw. Brakuje zatem uzasadnienia związku między uniwersalnością a prawem nauki. Taki związek jest szczególnie mało uchwytne na gruncie regularnościowej teorii praw: „jestem przekonany, że [...] uniwer-

salność nie jest warunkiem *sine qua non* praw, nawet nie «fundamentalnych» jako przeciwstawionych «wtórnym prawom»” (Earman 1978, 180).

Niezależnie więc od tego, czy uniwersalność jest jedną z właściwości praw nauki, nie stanowi ona poza projekcyjnością istotnego dodatku do roli eksplanacyjnej uogólnień. Załóżmy, że M' jest projekcją modelu przyczynowego M . Nietrywializująca analiza uniwersalności w sensie wyróżnionym przez Earmana może przebiegać na jeden z dwóch sposobów. Model M' może zostać uznany za uniwersalny w tym sensie, że jego dziedzina jest nieograniczona. Określmy ten sens uniwersalności jako *internalny*.

M' może być uniwersalny także w sensie *modularnym*. Mianowicie projektowana w M' zależność $F: \rightarrow G$ obowiązuje niezależnie od tego, w jakiej dziedzinie została zrealizowana czynność A prowadząca do pozyskania danych interwencyjnych dla tej zależności.

W sensie internalnym uniwersalność nie wiąże się w żaden sposób z rolą eksplanacyjną uogólnień. Powodem zasadniczym tego jest, iż tak sformułowana nie może generować niesprzecznych opisów interwencji prowadzących do pozyskania danych interwencyjnych dla projektowanej zależności. Taka interwencja wymagałaby zakwestionowania projekcji, dla której miała pozyskać dane.

W przypadku natomiast uniwersalności w sensie eksternalnym zachodzi wymagany związek z rolą eksplanacyjną uogólnień. Dla poszczególnych jednak przypadków wyjaśnień, gdy czynności pozyskiwania danych interwencyjnych dla $F \rightarrow G$ są realizowane w *konkretnej* dziedzinie, uniwersalność nie wprowadza nowego elementu w charakterystyce tego związku z rolą eksplanacyjną.

6.3 Projekcyjność uogólnień a informatywność i prostota

Za D. Lewisem jako kolejną z właściwości praw nauki przyjęło się uznawać parę: informatywność-prostota. Lewis wprowadził tę właściwość praw nauki jako ich istotną charakterystykę na gruncie przyjmowanej przez siebie regularnościowej koncepcji praw. Niezależnie jednak od tego, czy za teorię regularnościową uznamy, że odpowiedni układ tych własności determinuje istotę prawa nauki, należy uznać ten układ za ważny element każdej charakterystyki praw nauki.

Lewis pozostawia charakterystykę prostoty i informatywności praw w swoich późniejszych pracach niemal tak enigmatyczną, jak w pierwszym tekście, gdzie o nich mówi po raz pierwszy (Lewis 1973):

Jako hipotezę roboczą przyjmuję teorię praw utrzymywaną przez F. P. Ramseya w 1928 r.: prawa są „konsekwencjami tych sądów, które powinniśmy wybrać jako aksjomaty, jeśli wiedzielibyśmy wszystko i tę wiedzę uporządkowalibyśmy tak prosto, jak to możliwe w postaci systemu dedukcyjnego”. [...] Niezależnie od tego, czego się ostatecznie dowiemy, istnieją (jako przedmioty abstrakcyjne) niezliczone prawdziwe systemy dedukcyjne: dedukcyjnie zamknięte aksjomatyzowalne zbiory zdań prawdziwych. Spośród tych prawdziwych systemów niektóre dają się zaksjomatyzować *prościej* niż inne. Ponadto niektóre z nich mają większą *siłę* lub *zawartość informacyjną* od innych. (Lewis 1973, 73)

Istotnym elementem wprowadzonej przez Lewisa modyfikacji wyjściowej teorii Ramseya²³ jest kwestia zrównoważenia prostoty i informatywności praw jako cech przeciwstawnych:

Zalety prostoty i siły mają tendencję do konfliktu. Prostotę bez siły można znaleźć w czystej logice, siłę bez prostoty w (dedukcyjnie zamkniętym) almanachu. Niektóre systemy dedukcyjne nie są oczywiście ani proste, ani silne. W systemie dedukcyjnym cenimy właściwie zrównoważoną kombinację prostoty i siły – tak wiele obu, na ile pozwoli nam prawda oraz nasz sposób zrównoważenia. (Lewis 1973, 73)

Z uwagi na brak pełnej wiedzy o wszystkich prawdach (wszechwiedzy), nie możemy mówić o zbiorze wszystkich prawdziwych systemów dedukcyjnych. Pozostaje więc mówić o tych, które są najlepsze z uwagi na posiadaną przez nas wiedzę oraz kryterium zrównoważenia prostoty i informatywności.

Istotne uzupełnienie tej teorii, które Lewis dodał w 1983 (Lewis 1983/2000), dotyczy eksplikacji prostoty. Biorąc pod uwagę możliwość formułowania teorii dedukcyjnych w różnych językach i dopuszczając wśród nich języki z nowotworami, jak ziebieski, natrafiamy na poważny problem w eksplikacji kryterium prostoty. Bez żadnego ograniczenia dla języka, w którym formułowane mają być najlepsze teorie dedukcyjne, kryterium prostoty byłoby względne w stosunku do języka i w związku z tym nie otrzymalibyśmy żadnego kryterium dla praw naukowych.

Modyfikacja wprowadzona przez Lewisa miała zapobiec tej trudności. Aby wykluczyć języki z nowotworami językowymi, a przez to samo uwolnić kryterium prostoty od względności w stosunku do języka, Lewis wprowadził kategorię własności naturalnych. Tylko takie języki, w których predykaty odnoszą się do własności naturalnych (zielony i niebieski, ale nie ziebieski i nielony), są dozwolone przy formułowaniu prawdziwych teorii dedukcyj-

²³ Przeciwnie prace Ramseya odczytuje Czeżowski, który interpretuje je jako wyraz sceptycyzmu co do praw, uznającego je za reguły, a nie zdania; (por. 1947/1958,107).

nych, które następnie mają być zrównoważone pod względem prostoty i informatywności.²⁴

Wprowadzone przez Lewisa ograniczenie nałożone na wiedzę ogólną, a wymuszone odwołaniem do własności naturalnych w eksplikacji prostoty, jest zbyt restryktywne w stosunku do wyjaśnień dopuszczanych na gruncie tej wiedzy ogólnej. Jako ilustrację można przytoczyć dyskusję N. Swansona i C. Grangera (1997), w jaki sposób przejść od modelu ekonometrycznego w postaci opisanego przez C. Simsa (1980) do modelu przyczynowego (strukturalnego).

Takie przejście dokonuje się przez dekompozycję Choleksy'ego. Problem polega na tym, że istnieją alternatywne wersje takiej dekompozycji w zależności od sposobu uporządkowania zmiennych występujących w modelu.

Swanson i Granger odrzucają założenie wierności,²⁵ a w jego miejsce przyjmują strukturę łańcuchową uporządkowania zmiennych. W tym przypadku założenie wierności mówi, że nie ma innych zależności przyczynowych między zmiennymi niż zaobserwowane w danych (nie można przyjmować, że są między zmiennymi oddziaływania, których nie zaobserwowano w danych, bo się wzajemnie anulują).

Założenie wierności dopuszcza więc możliwość takiego uporządkowania między zmiennymi, w którym są liczne między nimi powiązania, a w szczególności uporządkowanie wedle struktury łańcuchowej. Przyjęte przez Swansona i Grangera założenie jest więc znacznie silniejsze niż założenie wierności, gdyż nie tylko eliminuje niezaobserwowane oddziaływania między zmiennymi, które się rzekomo anulowały, ale także eliminuje możliwość innych oddziaływań między zmiennymi niż przedstawione w strukturze łańcuchowej. Tym samym Swanson i Granger spośród możliwych wyjaśnień eliminują te wszystkie, w których struktura zależności między zmiennymi jest skomplikowana, ale zgodna z danymi, i pozostawiają jedynie zależności łańcuchowe.

Jeśli teraz przyjąć, że Swanson i Granger wprowadzają to ograniczające założenie, kierując się tym, że wyznacza ono „własności naturalne”, otrzy-

²⁴ Dyskusję Lewisa koncepcji praw naukowych w kontekście teorii regularnościowej podejmuję w kolejnym rozdziale. Tutaj natomiast omawiam wyłącznie kwestię związku między prostotą i informatywnością uogólnień a ich mocą eksplanacyjną. Szereg względów decyduje o tym, że dostrzegalny jest tu brak związku z mocą eksplanacyjną uogólnień, które są projekcyjne.

²⁵ Por. Punkt 4.4.

mujemy wyrazistą ilustrację tego, jak przyjęte przez Lewisa ograniczenie na wiedzę ogólną bezpodstawnie eliminuje niektóre z możliwych wyjaśnień. Przyjmując za Lewisem takie ograniczenie, dopuszcza się tę właśnie możliwość, że poza dopuszczonym przez nie możliwym wyjaśnieniem, znajdzie się faktyczne wyjaśnienie. Dlatego też można uznać, że w przypadku uogólnień projekcyjnych prostota i informatywność nie mają związku z mocą eksplanacyjną.

Inny poważny powód zakwestionowania tego związku podaje B. van Fraassen. Zwraca on uwagę, że proponowane przez Lewisa zrównoważenie prostoty i informatywności jest niezgodne z dynamiką nauki (1989, 55-59). Fakt, że naukowcy posługują się kryterium prostoty i informatywności względem dostępnej im *historycznie* wiedzy ogólnej oraz że być może posługują się także innymi kryteriami, gwarantuje postęp nauki – co możemy stwierdzić, patrząc na współczesną naukę i porównując ją z nowożytną, ale w żaden sposób nie gwarantuje osiągnięcia tego zrównoważenia, o którym pisze Lewis. Przywołując przykład Swansona i Grangera – nie mamy logicznej gwarancji, że dokonana dekompozycja Cholesky'ego oparta będzie na łańcuchowym, a nie innym, zgodnym z założeniem wierności, uporządkowaniu zmiennych.

Na koniec można krótko przywołać inne jeszcze zarzuty stawiane przez van Fraassena wobec związku między zrównoważeniem prostoty i informatywności a mocą eksplanacyjną. Po pierwsze, moc eksplanacyjna jest ściśle związana z informatywnością. Może się jednak okazać, że informatywność uogólnień występujących we wszystkich najlepszych systemach dedukcyjnych nie jest wystarczająca do tego, by podać wyjaśnienie.

Po drugie, nie jest wykluczone, że wyjaśnienie jest zależne od innej jeszcze cechy niż prostota i informatywność. W związku z tym uogólnienia najlepiej zrównoważone pod tymi dwoma względami mogą być bardzo źle zrównoważone pod względem tej dodatkowej cechy, która jest decydująca dla wyjaśniania. Po trzecie, umiejscowienie kryteriów prostoty i informatywności na końcu dziejów idealnej nauki pozostawia duże wątpliwości co do ich związku z faktycznie uprawianą nauką i podawanymi w niej wyjaśnieniami.

6.4 Projekcyjność uogólnień a konieczność praw przyrody

Inną tradycyjnie przypisywaną prawom naukowym właściwością jest ich konieczność. Filozofowie proponują różne jej rozumienia: od pojęciowej po

fizyczną.²⁶ Za van Fraassenem (1989, 28-30) przywołam podstawowe sposoby jej rozumienia. W pierwszym przypadku prawa przyrody wśród innych praw wyróżnia związek z prawdziwością, który krótko wyrażony przez wnioskowanie „Prawem przyrody jest, że P , a zatem P' ”. To wnioskowanie przestaje być prawomocne, jeśli w przesłance słowo „przyrody” zastąpić przez np. „społeczny”. Podobnie jeśli za „prawo przyrody” podstawić np. „dobrze potwierdzona i powszechnie akceptowana teoria”.

Ponadto wyrażenie „prawo przyrody” jest intensjonalne, gdyż nie dopuszcza dowolnego zastępowania go wyrażeniami równozakresowymi. Zmodyfikuję wcześniejszy przykład: „Prawdą jest, że każde zwierzę mające serce ma puls. Każde zwierzę mające serce ma nerkę. Zatem prawdą jest, że każde zwierzę mające nerkę ma puls”. Po podstawieniu w miejsce „prawdą” wyrażenia „prawem przyrody” otrzymamy wnioskowanie nieprawomocne.

Tym dwóm logicznym obserwacjom co do związku praw przyrody z koniecznością przeciwstawię dwa inne, gdzie konieczność jest cechą substancjalną praw przyrody. Pierwsza z nich ilustruje wnioskowanie „Prawem przyrody jest, że P , a zatem P jest konieczne”. Ten rodzaj konieczności van Fraassen określa jako konieczność przekazaną. Inne stosowane tu określenia to konieczność fizyczna lub nomologiczna.

Mocniejszy jeszcze sens konieczności to konieczność samych praw. Zgodnie z tym poglądem – rzadko reprezentowanym współcześnie – prawa przyrody nie mogłyby być inne nawet gdyby świat był inaczej urządzony. Z tej konieczności wynikają pozostałe.

Stawiając pytanie o rolę eksplanacyjną praw inną niż projekcyjnych uogólnień, odniosę się do wyróżnionej wyżej konieczności przekazanej w sensie van Fraassena, którą bardziej standardowo będę określał jako konieczność nomologiczną. Powstaje więc pytanie, czy konieczność nomologiczna stanowi istotny dodatek do roli uogólnień projekcyjnych.

Aby odpowiedzieć na to pytanie, należałoby podać dokładniejszą charakterystykę konieczności nomologicznej. W tym jednak punkcie natrafiamy na zasadniczą trudność, którą stwierdził R. Pargetter (1984), a standardowo określa się mianem problemu identyfikacji.²⁷ Zasadniczo problem ten polega

²⁶ Por. (van Fraassen 1977b; Amsterdamski 1983, 75-83; Kistler 2002). D. Szejnberg dokonała szczegółowego przeglądu najważniejszych historycznie interpretacji konieczności praw przyczynowych (1931/1990).

²⁷ Por. (van Fraassen 1989, 72-78). Problem identyfikacji jest problemem redukcyjnych koncepcji praw przyrody, które podają treścią ich charakterystykę. Dyskusję podejść nie-

na tym, w jaki sposób – unikając kołowości i odniesienia do praw przyrody – podać charakteryzację konieczności nomologicznej.

Przyjmę, że zbiór światów rzeczywiście możliwych jest ścisłym podzbiorem wszystkich logicznie możliwych. Między rzeczywiście możliwymi światami zachodzi relacja dostępności taka, że prawo przyrody, które jest konieczne w jednym ze światów rzeczywiście możliwych, jest także prawdziwe we wszystkich pozostałych dostępnych z tego świata. Pytanie o identyfikację konieczności nomologicznej sprowadza się więc do pytania o taką charakterystykę relacji dostępności między światami, która nie odwoływałaby się do kategorii prawa przyrody.

Podanie nietrywialnej charakterystyki takiej relacji (która nie byłaby relacją między wszystkimi światami albo żadnym) nie wyodrębni jednej relacji. Dla danej charakterystyki można bowiem znaleźć izomorficzną relację, która spełnia dokładnie te same warunki (van Fraassen 1989, 72).

Nie może powieść się również próba określenia tej relacji przez odniesienie do historii światów. Są dwa możliwe sposoby określenia superwencji konieczności nomologicznej na historii światów możliwych. W pierwszym wybiera się pewien moment czasu t i określa jako nomologicznie konieczne takie P , które jest prawdziwe we wszystkich światach mających do czasu t tę samą historię. Jednak to, jaka była ta historia do czasu t , wcale nie musi być wyłącznie wypadkową *tylko* praw przyrody, gdyż może też zależeć od pewnych zdarzeń losowych.

W drugiej – konieczność nomologiczna superwenuje na wspólnej historii światów możliwych *w każdym* momencie czasowym poprzedzającym wybrany t . Otrzymuje się więc nieintuicyjny wniosek, że o tym, co jest konieczne nomologicznie i co jest prawem przyrody, decyduje np. trwająca 2 nanosekundy historia światów możliwych. Może być, że P nie jest prawem w tych światach możliwych, które przez 2 pierwsze nanosekundy dzielą wspólną historię, ale *mają inne prawa*. Inaczej mówiąc, w tym przypadku otrzymalibyśmy zbyt mocną zależność konieczności nomologicznej oraz prawa dla zdarzeń, jakie zaszły w danym świecie.

Nie mając zatem drogi do precyzyjniejszego określenia konieczności nomologicznej, a tym samym konieczności praw przyrody, pozostają przy projekcyjności uogólnień jako tej cesze, która decyduje o ich roli eksplanacyjnej.

redukcyjnych przesuwam do kolejnego rozdziału, gdyż charakteryzują one prawa przyrody wyłącznie ekstensjonalnie przez ich role funkcjonalne (Psillos 2002, 204).

Projekcyjność modeli przyczynowych

W poprzednim rozdziale, analizując właściwości praw przyrody, wskazałem, że dla pełnienia roli eksplanacyjnej, zwłaszcza w badaniach nieeksperymentalnych, którą tradycyjnie zarezerwowano wyłącznie dla praw, wystarczające jest posiadanie przez uogólnienia cechy projekcyjności. Przechodząc do analizy różnych koncepcji praw, będę zmierzał do wykazania, że projekcyjność jest także warunkiem koniecznym zdeterminowania roli eksplanacyjnej praw. Zasadnicza strategia przy argumentacji za tą tezą polega na wykazaniu, że uogólnienia projekcyjne unikają fundamentalnych trudności redukcyjnych koncepcji praw przyrody, a mianowicie problemu identyfikacji i inferencji (Pargetter 1984; van Fraassen 1989; Psillos 2002, 165).

Problem identyfikacji to jeden z tradycyjnych problemów teorii praw przyrody. Dotyczy on wyodrębnienia praw przyrody wśród uogólnień innego rodzaju, zwłaszcza tzw. uogólnień przypadkowych. Większość tych ostatnich zawiera predykaty pozycyjne, odnoszące się do konkretnych lokalizacji lub czasu, np. wszystkie monety w mojej lewej kieszeni to pięciogroszówki. Jednak niektóre z przypadkowych uogólnień w formie syntaktycznej i semantycznej bliskie są uogólnieniom *prawopodobnym*. I tak uogólnienie „Wszystkie kule ze złota mają średnicę mniejszą niż kilometr” uznaje się za przypadkowe, bo nic (wśród znanych praw przyrody) nie wyklucza możliwości istnienia takiej kuli. Podobne semantycznie uogólnienie „Wszystkie kule z uranu mają średnicę mniejszą niż kilometr” nie jest jednak przypadkowe, gdyż jest konsekwencją znanych praw fizyki.

Poniżej omawiam trzy podstawowe strategie w konstruowaniu koncepcji praw przyrody. Zgodnie z koncepcją regularnościową, prawa to takie uogólnienia, które są wyodrębnione przez najlepszą dedukcyjną systematyzację

wiedzy osiągniętej przez naukę idealną. Jak pokazuję w Punkcie 7.1, takich systematyzacji może być wiele, dlatego problem identyfikacji nie ma tu jednoznacznego rozwiązania.

Kolejną koncepcją jest koncepcja koniecznościowa, zgodnie z którą prawa przyrody są konieczne.¹ Jak częściowo pokazałem w Punkcie 6.4 oraz pełniej w Punkcie 7.2 poniżej, brak jest dostatecznej charakterystyki konieczności, by można było uznać ją za podstawę rozwiązania problemu identyfikacji. Charakterystyka poprzez formalne własności relacji między światami możliwymi czy w kategoriach superweniencji na historii światów możliwych nie daje zadowalającej odpowiedzi.

Problem identyfikacji stanowi zasadniczą trudność dla nieredukcyjnej koncepcji praw przyrody. Tutaj zakłada się bowiem, że prawa przyrody są nieredukowalne do innych kategorii (jak regularność czy konieczność), a ich charakterystyka podawana jest poprzez ich rolę funkcjonalną w nauce. Jak pokażę w Punkcie 7.5, trudno jednak podać taką nieredukcyjną charakterystykę, która nie byłaby – choćby pośrednio – kołowata.

Analizując redukcyjne koncepcje praw przyrody, wracam do tezy, że ograniczenie się w danej koncepcji do uogólnień projekcyjnych prowadzi do rozwiązania problemu identyfikacji, a także problemu inferencji.

Jeśli schematycznie prawo przyrody reprezentuje formuła „Prawem przyrody jest, że P ”, to problem inferencji dotyczy tego, jak uzasadnić, że z takiej formuły wynika P , które odnosi się do obserwowanych w badaniach naukowych regularności w świecie. Na gruncie teorii regularnościowej można sformułować pewne rozwiązanie tego problemu, ale – jeśli prześledzić to dokładniej – powstaje pytanie, czy te implikowane regularności są zbieżne z regularnościami stwierdzanymi w badaniach naukowych.

Dużą trudność problem inferencji stanowi natomiast dla koncepcji koniecznościowej. Przyjęcie, że konieczność praw przyrody ma charakter pierwotny i nieeksplikowalny, prowadzi do impasu w podaniu rozwiązania dla problemu inferencji. Z kolei eksplikacja kategorii konieczności prowadzi do trudności w podaniu jednoznacznego rozwiązania dla problemu identyfikacji.

W konstrukcji nieredukcyjnych koncepcji praw przyrody wiele uwagi poświęca się problemowi inferencji. Pozostaje jednak pytanie, czy można takie rozwiązanie przeformułować w taki sposób, aby umożliwić takie niekołowate rozwiązanie problemu identyfikacji.

¹ Por. (Such 1967, 32-34).

Jak wynika z powyższych uwag, podanie zadowalającej koncepcji praw przyrody wymaga pewnego zrównoważenia rozwiązań problemów identyfikacji i inferencji. W odniesieniu do wszystkich z analizowanych koncepcji, podkreślam poniżej, że takie zrównoważenie jest osiągalne tylko dla uogólnień projekcyjnych. Stąd też na końcu zostanie wyprowadzony ogólny wniosek, że rolę eksplanacyjną w badaniach naukowych należy przypisać nie prawom przyrody, a uogólnieniom projekcyjnym.

7.1 Uogólnienia projekcyjne a regularnościowa koncepcja praw

Zasadniczym rysem tej koncepcji jest to, że prawa przyrody są pewnego rodzaju regularnościami oraz że nie ma podstaw ontologicznych dla wyróżnienia tych regularności spośród innych. Różnice stanowisk w tym podejściu dotyczą sposobu charakteryzowania różnicy między prawami przyrody a regularnościami innego rodzaju, zwłaszcza przygodnymi.

Tutaj skupię się na analizie stanowiska, które jest najpoważniejszą ze współczesnych koncepcji regularnościowych. Określa się je mianem najlepszego systemu. Zainicjował je J. S. Mill w *Systemie logiki*, określając, że prawa przyrody to te spośród regularności, na podstawie których można wyprowadzić inne (1843/1962 t. 2, s. 56-57).² Ramsey (1925) uzupełnił to określenie, dodając, że chodzi o dedukcyjną systematyzację naszej wiedzy w momencie, kiedy nauka byłaby już zakończona. Wśród możliwych kryteriów wyboru takiej systematyzacji dedukcyjnej są prostota oraz moc (zawartość informacyjna).

Lewis, najwybitniejszy współczesny kontynuator Ramsey'a, zwrócił uwagę na opozycyjność tych cech: zwiększanie prostoty zubaża moc, a zwiększanie zawartości informacyjnej danej systematyzacji wprowadza dodatkowe jej skomplikowanie. Z uwagi na brak wszechwiedzy nie będziemy wiedzieć jednoznacznie, która z dedukcyjnych systematyzacji jest najlepsza. Dlatego też zaproponował on dalszą modyfikację tej koncepcji: prawa przyrody to regularności, które występują we wszystkich najlepszych systematyzacjach dedukcyjnych osiągniętych przez w pełni rozwiniętą naukę.

Ponadto zastosowanie kryterium prostoty wymaga, jak wyjaśniałem w Punkcie 6.3, wprowadzenia kategorii własności naturalnych. Nie można bowiem ocenić prostoty, porównując systematyzacje wiedzy sformułowane w języku własności naturalnych oraz nienaturalnych.

² Por. (Zamecki 1967, 15-24).

Ostatecznie więc o tym, czy dana regularność jest prawem przyrody, przesądza przy tym podejściu pięć kryteriów: sformułowanie w odpowiednim języku (pozwalającym wyrazić własności naturalne), prawdziwość, prostota, moc oraz odpowiednie zrównoważenie dwóch ostatnich. Czy wystarczają one do tego, by uniknąć problemu identyfikacji i inferencji?

Zacznę od problemu inferencji. Konstrukcja praw przyrody w koncepcji najlepszego systemu gwarantuje, że znajdują się one wśród podstawowych twierdzeń systematyzacji dedukcyjnej u kresu nauki. Można więc zatem będzie wyprowadzić z nich wszystkie pozostałe regularności – na mocy tego, czym ma być najlepszy system.

Przed przejściem do problemu identyfikacji zwrócę uwagę na pewną subtelność w tym rozwiązaniu. Wyprowadzenie z praw przyrody wszystkich pozostałych regularności może być rozumiane warunkowo bądź bezwarunkowo. W sensie warunkowym będą to wszystkie regularności, dla których najlepszy system podaje systematyzację dedukcyjną. I w tym sensie problem inferencji byłby rozwiązany. Z samej konstrukcji stanowiska najlepszego systemu wynika, że zawsze prawomocne jest wnioskowanie „Prawem przyrody jest, że P , a zatem P' ”.

W sensie bezwarunkowym chodzi o wyprowadzanie „pozostałych” regularności, które są stwierdzane w nauce. To, czy w tym przypadku koncepcja najlepszego systemu podaje rozwiązanie problemu inferencji, zależy już będzie od tego, jaka jest relacja między klasą regularności, dla których podawana jest systematyzacja dedukcyjna przez naukę w pełni rozwiniętą, a klasą regularności podawaną przez faktycznie uprawianą naukę. Czy te dwie klasy się pokrywają, zależy z kolei od rozwiązania drugiego problemu, a mianowicie problemu identyfikacji.

Jest wiele przesłanek ku temu, by podejrzewać, że te dwie klasy regularności się różnią i że problem identyfikacji stanowi poważną trudność dla koncepcji najlepszego systemu. Naświetlę tutaj dwa zasadnicze.³

Żądanie używania przez naukowców filozoficznie usankcjonowanego języka (własności naturalnych) wyprowadza nas poza klasę regularności stwierdzanych w nauce. Jeśli przyjąć, że istnieje niezależna od postępowania naukowców (nieretrospektywna) charakterystyka takiego języka, wówczas bezzasadnie ograniczałaby postępowanie naukowców.

Jedyne bardziej specyficzne wskazanie takiego języka, jakie może się tu nasuwać, to wynikające z ewolucji. Selekcja w ewolucji ukierunkowała

³ Rozwiniętą argumentację podaje także van Fraassen (1989, 51-64).

człowieka na rozpoznawanie własności naturalnych (zagrożenie np. trucizną czy dobieranie partnerów) i odrzucanie nienaturalnych. Jednak gwarantujące przeżycie i rozmnażanie własności rozpoznawane przez człowieka nie mają bezpośredniego przełożenia na predykaty występujące w nauce, np. kolor czerwony jest istotnym sygnałem, ale w fizyce jest pochodną innych własności. Nawet więc jeśli naukowcy rozpoznają klasy naturalne oraz jeśli predykaty w nauce wyodrębniają klasy naturalne, to z tego, co powiedziano, nie wynika, że w nauce pozostają teorie sformułowane we właściwym języku („naukowych” klas naturalnych), które wyodrębniono ze względu na własności naturalne jako rozpoznawalne przez człowieka niezależnie od nauki.

Jako kryterium retrospektywne, żądanie sformułowania w kategoriach naturalnych własności jest dla koncepcji najlepszego systemu bezużyteczne. Nie wnosłoby bowiem żadnej treści ponad to, że naukowiec na danym etapie badań faktycznie dokonał wyboru jednego z możliwych modeli.

Pozostaje zatem możliwość, że konfrontacja uogólnień z danymi prowadzi do wyłonienia tych, które są sformułowane w kategoriach własności naturalnych. W tym jednak przypadku jedyny dowód na to, że uogólnienie jest sformułowane w poprawnym języku własności naturalnych, może pochodzić z faktu, że jest ono w gronie tych, które przetrwały wszystkie próby konfrontacji z doświadczeniem i jest w najlepszym systemie u kresu nauki. Jako takie to kryterium byłoby czysto retrospektywne i jak podkreśliłem – bezużyteczne dla tej koncepcji.

Powyższe rozważania prowadziłyby więc do trywializacji własności naturalnych na gruncie koncepcji najlepszego systemu bądź też do uzależnienia ich od faktycznej historii nauki i wszystkich wpływających na nią przypadków, np. tego, jak rozwinię się nauka przy braku katastrof zagrażających populacji naukowców oraz przy zaistnieniu takich katastrof. Trywializacja jednak nie jest tam dopuszczana za cenę – przypomnijmy – relatywizacji kryterium prostoty. Brak wyodrębnienia naturalnych własności od nienaturalnych blokuje całe przedsięwzięcie wyodrębnienia prostych, a tym samym najlepszych systematyzacji dedukcyjnych.

Z kolei uzależnienie rozumienia własności naturalnych, a więc i praw przyrody, od przypadkowych zdarzeń i możliwego rozwoju wypadków, wpływa na zakwestionowanie celowości całego przedsięwzięcia, które miało je dopiero ugruntować.

Drugi powód do podejrzenia, że prawa przyrody wyodrębnione przez naukę u kresu w sposób opisywany przez koncepcję najlepszego systemu, to uproszczona charakterystyka dynamiki nauki. Po pierwsze, prostota i moc

dla Lewisa to kryteria syntaktyczne i semantyczne, podczas gdy w rzeczywistości uprawianej nauce są to kryteria pragmatyczne. Zgodna z koncepcją najlepszego systemu ocena prostoty wymagałaby przełożenia uogólnień zaproponowanych w badaniach na język nauki u kresu, w którym można by je porównać ze sobą i podać jednoznaczną odpowiedź. Taka możliwość nie jest dostępna jednak dla naukowców, którzy dokonują oceny prostoty w takim (niedoskonałym) języku, w jakim są sformułowane na obecnym etapie rozwoju nauki. Taka ocena prostoty jest więc uwarunkowana historycznie, a prawdopodobnie także obciążona subiektywną oceną z punktu widzenia aktualnie przyjmowanych przez naukowców wartości.

Podobnie z oceną mocy uogólnień. Nawet jeśli przyjąć, że istnieje kryterium informatywności, które daje obiektywną ocenę w tak złożonym kontekście, jakim są uogólnienia proponowane w nauce, to i tak w faktycznie podejmowanych decyzjach w nauce stosowane jest inne kryterium. Tym kryterium jest zawieranie wartościowych czy ważnych informacji, co z kolei ocenia się względem aktualnie stawianych pytań.

I prostota, i moc uogólnień są więc oceniane w nauce względem danej sytuacji historycznej, a nie wobec logicznie prostych i obiektywnych kryteriów prostoty i mocy znanych u kresu nauki. Stąd rozbieżność między zbiorem uogólnień, do którego zmierza faktycznie uprawiana nauka, a tym, który wyodrębniłaby koncepcja najlepszego systemu.

Po drugie, oprócz uwzględnionych kryteriów w koncepcji najlepszego systemu mogą istnieć również inne. Dopuszczenie jednak takiej możliwości całkowicie zmienia obraz nauki u kresu badań. Załóżmy bowiem, że wybieramy zgodnie z określonym zrównoważeniem cech prostoty i mocy, ale warunkowo względem innego kryterium. Wówczas nie mamy gwarancji, że to inne kryterium – np. zgodności z innymi uogólnieniami poza danym kontekstem badawczym⁴ – nie wyeliminowało uogólnień najprostszych i najmocniejszych. Osiągnięte zatem zrównoważenie prostoty i mocy w grupie uogólnień spełniających dodatkowe kryteria lokowałoby nas poza tym zbiorem uogólnień, które zostałyby bezwarunkowo wyłonione przez zrównoważenie prostoty i mocy.

Po trzecie, zwrócenie uwagi na punkt wyjścia nauki zmienia obraz nauki proponowany przez Lewisa. Istotnym elementem w historycznej trajektorii rozwoju nauki jest jej punkt wyjścia – przyjęcie określonej kategoryzacji świata, wyodrębnienie podstawowych zależności między nimi itp. Nauka u

⁴ Przykładem może być odrzucenie przez W. T. Kelvina teorii ewolucji jako wymagającej znacznie dłuższego wieku Ziemi niż ten znany ówczesnie z fizyki (van Fraassen 1989, 58).

kresu pozostanie funkcją tego punktu wyjścia, nawet jeśli w trakcie całego tego procesu dokonywano by wymaganego zrównoważenia prostoty i mocy. Nie ma gwarancji, że przyjęcie innego punktu wyjścia doprowadziłoby przy zastosowaniu tego samego zrównoważenia prostoty i mocy do otrzymania zupełnie rozbieżnego zbioru najlepszych dedukcyjnie systematyzacji. Tym samym do otrzymania praw przyrody, które byłyby prostsze i mocniejsze (bądź nie miałyby nic wspólnego) od praw przyrody otrzymanych w rozwoju nauki z jej faktycznym punktem wyjścia.

W końcowej części podam powody, dla których można sądzić, że wskazane tu problemy nie miałyby miejsca, gdyby w miejsce praw przyrody wprowadzić uogólnienia projekcyjne. Przypomnę, że za projekcyjne uznajemy takie uogólnienie, które odnosi stwierdzoną w danej dziedzinie zależność przyczynową między X a Y na nową dziedzinę w taki sposób, że można określić, jakie czynności pozwolą na uzyskanie danych pozwalających na stwierdzenie, czy jest to też zależność przyczynowa w tej nowej dziedzinie.

Podobnie więc jak w przypadku praw przyrody, problem inferencji jest warunkowo rozwiązany na mocy postulatu. Z definicji projekcyjności wynika bowiem, że $X : \rightarrow Y$ implikuje $X \rightarrow Y$ dla dziedziny, w której obowiązuje ta zależność przyczynowa, a także dla nowej – projekcyjnej – dziedziny, pod warunkiem uzyskania odpowiednich danych oraz spełnienia kryterium proceduralnego.

Dlaczego więc nie powstaje rozbieżność między uogólnieniami projekcyjnymi wyodrębnionymi na podstawie kryteriów koncepcji najlepszego systemu a osiągniętymi w faktycznie uprawianej nauce?

Wcześniej podane argumenty odwoływały się do dwóch racji: wymogu formułowania praw przyrody w kategoriach własności naturalnych oraz wy-preparowanego semantycznie opisu dynamiki nauki. Zacznę więc od kwestii pierwszej. Uogólnienia projekcyjne są formułowane względem wcześniej przyjętych modeli przyczynowych. Nie przesądza to jednak o tym, jaki jest ich sposób sformułowania: w kategoriach naturalnych czy nienaturalnych. Pewnym ograniczeniem jest wprowadzony w przyjętym tu rozumieniu projekcyjności wymóg podania czynności prowadzącej do uzyskania danych interwencyjnych dla projektowanej zależności. Nawet jeśli przyjąć, że przyjęte modele przyczynowe bazują na naturalnej kategoryzacji przedmiotów, to ze względu jednak na daleką specjalizację czynności badawczych, zwłaszcza wykorzystujących aparaturę wyspecjalizowaną badawczą, nie można oczekiwać – jak zaznaczyłem to wcześniej – odpowiedniości między tym, co byłoby naturalne na poziomie potocznym i wyspecjalizowanym. Na pewno

taka odpowiedniość nie jest konieczna do tego, by sformułować wymagany opis.

Przejdźmy więc do drugiego źródła rozbieżności między nauką u kresu wedle koncepcji najlepszego systemu a rzeczywistą nauką. Proponowane rozumienie projekcyjności pozostawia jedynie pewną wartość zwrócenia uwagi możliwość rozbieżności między faktyczną a idealną nauką u kresu pod względem kryterium mocy. Taka rozbieżność zaistniałaby wówczas, gdyby po dokonaniu projekcji z dziedziny D na dziedzinę D' możliwe było dokonanie projekcji na dziedzinę D'' , ale nie w odwrotnej kolejności. A więc dokonanie projekcji z D na D'' w jakiś sposób uniemożliwiłoby dokonanie kolejnej projekcji na D' . Gdyby z jakiegoś powodu istniała taka asymetria między D' a D'' , wówczas wybranie w historii nauki D'' mogłoby skutkować rozbieżnością między uzyskanymi faktycznie uogólnieniami projekcyjnymi a zalecanymi zgodnie ze zrównoważeniem kryterium prostoty i mocy.

Jedynym przypadkiem, kiedy z punktu widzenia projekcyjności mogłaby powstać taka asymetria, jest sytuacja związana z pozyskiwaniem danych interwencyjnych. Załóżmy, że między dziedzinami D' a D'' istnieje pewna asymetria, np. czasowa lub genetyczna. Oznaczmy ją przez „ \prec ”: $D' \prec D''$. Taka asymetria przekształcić się może w asymetrię między projekcjami z D' oraz D'' tylko wówczas, gdy opis czynności pozyskiwania danych interwencyjnych odnosi się do czynności A w danej dziedzinie: $A(D')$, która powoduje, że wcześniejsza asymetria między dziedzinami zostaje odwrócona: $A(D'') \prec A(D')$. Otrzymalibyśmy więc przypadek, gdy dokonanie projekcji z D na D' bezpowrotnie odbierałoby możliwość dokonania projekcji na D'' , a tym samym powstaje możliwość rozbieżności między zbiorem uogólnień uzyskanej przez naukę faktyczną a przez idealną naukę u kresu.

Czynność A w faktycznej nauce może być jednak zrealizowana dla pewnej skończonej poddziedziny D' . Otrzymalibyśmy zatem, po faktycznym zrealizowaniu A , partycję D' oraz D'' w taki sposób, że dla odnośnej poddziedziny D' i D'' zachodziłoby: $A(D'') \prec A(D')$, a dla pozostałych części tych dziedzin $D' \prec D''$. Przywrócona zostaje tym samym możliwość natrafienia na tę drugą projekcję w trakcie rozwoju nauki faktycznej, niezależnie od przygodnej historycznie kolejności wyboru dziedzin D' oraz D'' w dokonywaniu projekcji z dziedziny D .

Zastosowanie kryteriów selekcji uogólnień wskazanych w koncepcji najlepszego systemu do uogólnień projekcyjnych, w miejsce praw przyrody, nie prowadzi – jak pokazałem – do rozbieżności między nauką faktyczną a

idealną nauką u kresu. Wprowadzone określenie projekcyjności, z jednej strony, nie pretenduje do sterowania za pomocą uprzednich filozoficznych rozstrzygnięć (jak własności naturalne) dynamiką nauki. Z drugiej strony natomiast, bardziej niż pojęcie prawa zawęża logicznie możliwe przejścia na kolejnych etapach, by pozostać bliżej faktycznej dynamiki nauki.

7.2 Uogólnienia projekcyjne a koniecznościowa koncepcja praw

Regularności obserwowane w świecie to tylko przejaw praw przyrody, które są koniecznym związkiem własności przedmiotów. Taka ogólna charakterystyka stanowiska koniecznościowego wymaga dopełnienia poprzez dokładniejsze określenie natury tej konieczności oraz rodzaju własności przedmiotów, które wchodzą w taki konieczny związek. Dopiero wówczas można pytać, czy rozwiązany jest na gruncie tej koncepcji w sposób zadowalający problem identyfikacji oraz inferencji.

W. Kneale (1949), podając jako ilustrację takiej konieczności niemożność połączenia cech czerwieni i zieleni, na ponad dwadzieścia lat zasugerował, że jest to konieczność logiczna lub pojęciowa: Konieczne jest, że Z jest prawdziwe w świecie w wtw Z jest prawdziwe we wszystkich światach możliwych. Po dodaniu, że jako taka konieczność jest poznawalna tylko apriorycznie, została uznana za nieadekwatną do wyjaśnienia natury praw przyrody, które są poznawane aposteriorycznie:⁵

Konieczność jest kategorią logiczną, której nie da się zastosować jednoznacznie do rzeczywistości fizycznej. (Mazierski 1995, 66)

Dopiero argumentacja S. Kripkego (1972/1988) utarowała drogę nowym koncepcjom koniecznościowym, które przyjmowały, że jest to konieczność przygodna (np. obowiązująca tylko w tych światach możliwych, w których występują dane własności) oraz poznawalna aposteriorycznie: Konieczne jest, że Z jest prawdziwe wtw Z jest prawdziwe w świecie w wtw Z jest prawdziwe we wszystkich światach dostępnych z w .

Problemy nieredukcyjnej analizy praw w kategoriach modalnych Pargettera omówiłem w Punkcie 6.4. Pokazałem tam również, w jaki sposób problemy identyfikacji i inferencji ujawniają braki redukcyjnej analizy praw przyrody jako superwenujących na historii światów możliwych.

⁵ Znacznie wcześniej koncepcję koniecznościową przyczynowości na gruncie polskim wyeksplikował J. Łukasiewicz (1907/1961); por. (Bigaj 1994; Mazierski 1995, 64-65).

Zacznę od teorii zaproponowanej przez Tooleya (1977). Określił on hierarchię, którą tworzą partykularia (osoby, przedmioty) oraz uniwersalia. Partykularia są na poziomie zerowym, a ich własności i relacje między nimi na poziomie pierwszym. Ogólnie, uniwersale, jakim jest własność lub relacja, jest na poziomie o jeden wyższym niż najwyższy poziom bytów, którym przysługuje. Relacja ukonieczniania R jest przygodna (są przedmioty, między którymi może zachodzić, i są takie, między którymi może nie zachodzić), nieredukowalna do poziomu niższego niż drugi, a z twierdzenia, że R zachodzi między uniwersaliami poziomu pierwszego, wynika, że określona inna relacja zachodzi pomiędzy egzemplifikacjami tych uniwersaliów. W przypadku, gdy wszystkie egzemplifikacje pierwszego uniwersale są egzemplifikacjami drugiego, mówi się o relacji ukonieczniania w sensie właściwym.

W ten sposób Tooley rozwiązuje problem identyfikacji, podając charakterystykę relacji ukonieczniania nieredukowalną do innych własności. W szczególności nie jest ona redukowalna do charakterystyki bycia podzbiorem: każda egzemplifikacja F jest G .

Powstaje pytanie o rozwiązanie problemu inferencji. Nieredukowalność relacji ukonieczniania implikuje, że nie może wystąpić w niej żadna charakterystyka z poziomu poniżej drugiego. Jednak bycie podzbiorem (wszystkie przypadki F są przypadkami G) nie zawiera żadnych uniwersaliów z poziomu drugiego. Innymi słowy, podanie charakterystyki z poziomu drugiego w żaden sposób nie przesądza o tym, co dzieje się na poziomie niższym.

Van Fraassen, dyskutując koncepcję Tooleya, wprowadza następującą modyfikację, która mogłaby rozwiązać problem inferencji (1989, 101). Załóżmy, że istnieje relacja R' z poziomu trzeciego, która zachodzi między R a F oraz między R a G . Dla dowolnego X , które wchodzi w relację R z Y , relacja R' zachodzi między R a X i Y . Z tego można wyprowadzić wniosek, że wszystkie F są G .

Aby ustalić jednak, że mamy tutaj do czynienia z wynikaniem, należałoby z kolei przyjąć istnienie relacji R'' z poziomu czwartego, która – podobnie jak R' w stosunku do R – zagwarantowałaby, że to wynikanie faktycznie zachodzi. Nawet jeśli przyjąć taki regres, to z pewnością tak rozumiane prawa nie mogłyby służyć jako podstawa wyjaśniania, gdyż wymagałoby to zatrzymania tego regresu w pewnym punkcie.

Obecnie przejdę do dyskusji redukcyjnej analizy praw przyrody w kategoriach uniwersaliów. Przeanalizuję najbardziej systematycznie opracowaną koncepcję D. Armstronga (1978; 1983; 1993; 1997).

Jej zasadniczym elementem jest twierdzenie, że prawa przyrody są koniecznościową relacją między dwoma uniwersaliami, które są egzemplifikowane w partykulariach: bądź (niezdolnych do samodzielnego istnienia) stanach rzeczy, bądź w przedmiotach (zdolnych do samodzielnego istnienia). Symbolicznie tę relację zapisuje się jako $N(F, G)$, co oznacza, że uniwersale F ukoniecznia uniwersale G . Innymi słowy, każdy przedmiot egzemplifikujący F musi egzemplifikować także G .

W koncepcji Armstronga, dla uniknięcia kołowości przyjmuje się, że uniwersalia nie istnieją poza swoimi egzemplifikacjami. Gdyby nie uznać tej tezy, wówczas konieczne byłoby wprowadzenie relacji, np. R , która określałaby kiedy uniwersale F jest egzemplifikowane w jakimś partykularne.

Relacja ukonieczniania $N(-, -)$ nie jest jednak wolna od wieloznaczności (van Fraassen 1989, 105-106). Zasadniczo jest to relacja zachodząca między stanami rzeczy, gdzie bycie F danego partykularne a ukoniecznia jego bycie G : $N_1(a \text{ bycie } F, a \text{ bycie } G)$. Z zachodzenia relacji ukonieczniania wynika prawdziwość obu stanów rzeczy, a w konsekwencji prawdziwość stanu rzeczy, że a jest F oraz G .

Szczególnym rodzajem relacji N_1 jest relacja ukonieczniania ze względu na relację zachodzącą między F a G : $N_1(F, G)(a \text{ bycie } F, a \text{ bycie } G)$. W tym przypadku dodaje się informację, że bycie F danego partykularne a ukoniecznia jego bycie G tylko z uwagi na relację, jaka zachodzi między F a G . Ta relacja to również relacja ukonieczniania, ale zachodząca wyłącznie między uniwersaliami: $N_2(F, G)$. Zachodzenie tej relacji nie jest zdeterminowane przez żadne partykularia i dlatego Armstrong w *What is a law of nature?* (1983) traktuje ją jako pierwotną.

Powstaje więc pytanie, czy z tej relacji ukonieczniania wynika, że zbiór przedmiotów F jest podzbiorem G . Armstrong proponuje następujące rozwiązanie. Utożsamia uniwersale, którym jest relacja ukonieczniania między stanami rzeczy N_1 , ze stanem rzeczy, którym jest relacja ukonieczniania między uniwersaliami N_2 , twierdząc, że $N_1(F, G)(a \text{ bycie } F, a \text{ bycie } G)$ jest egzemplifikacją relacji $N_2(F, G)$. W efekcie więc zanika różnica między N_1 a N_2 . Pozostaje jednak pytanie, czy $N(F, G)$ implikuje, że $Fb \rightarrow Gb$.

Ponieważ, jak stwierdziłem wcześniej, relacja ukonieczniania zachodzi między stanem rzeczy a jest F oraz a jest G niezależnie od tego, jakie jest a , a jedynie w zależności od relacji N między uniwersaliami F i G , można więc wywnioskować, że z tego, iż b jest F oraz G , wynika, że $N(F, G)(b \text{ jest } F, b \text{ jest } G)$. Jeśli więc przyjąć, że relacja $N(F, G)$ jest realna (jest prawem przy-

rody), to każde wystąpienie łączne cech F i G dokonuje się nieprzypadkowo, ale ze względu na to prawo.

To jednak nie jest wniosek w takiej postaci, jaka jest potrzebna do rozwiązania problemu inferencji. Jak podkreśla van Fraassen, nie ma możliwości rozwiązania tego problemu. Bowiem między relacjami wśród uniwersaliów a relacjami wśród partykulariów nie zachodzi żadna relacja logiczna.

Czy można zatem uniknąć trudności wynikających z problemu inferencji na gruncie teorii koniecznościowej, przypisując rolę eksplanacyjną uogólnieniom projekcyjnym w miejsce praw przyrody?

Przypomnę, że za projekcyjne uznajemy takie uogólnienie, które odnosi stwierdzoną w danej dziedzinie zależność przyczynową między X a Y na nową dziedzinę w taki sposób, że można określić, dzięki jakim czynnościom uzyskamy dane pozwalające na stwierdzenie, czy jest to także zależność przyczynowa w tej nowej dziedzinie.

Zasadnicza idea rozwiązania problemu inferencji dla uogólnień projekcyjnych na gruncie koncepcji koniecznościowej związana z proponowanym tu podejściem proceduralnym sprowadza się do odpowiedniej identyfikacji ekstensji tych uogólnień.

Próba identyfikacji ekstensji uogólnień projekcyjnych poprzez frontalną eksplikację kategorii przyczynowości w kategoriach praw przyrody nieuchronnie doprowadzi nas do problemu regresu i braku przejścia inferencyjnego do zdań niemodalnych, jakie można zauważyć w przypadku koncepcji Tooleya oraz Armstronga. Analogicznie do relacji ukonieczniania, otrzymalibyśmy modalne określenie: to, że a jest F , jest przyczyną, że a jest G , na mocy prawa przyczynowego łączącego te dwa egzemplifikowane przez a uniwersalia. Nawet będąc w stanie pokazać, że w przypadku realizacji F oraz G w dowolnym przedmiocie, dokonuje się ona na mocy odnośnego prawa przyczynowego, nie moglibyśmy więc pokazać, że takie prawo obowiązuje w pełnej ogólności, zgodnie z wymaganą kwalifikacją modalną.

Proponuję więc określenie ekstensji uogólnień projekcyjnych na gruncie koncepcji koniecznościowej, które opierałoby się na wprowadzonym w części drugiej kryterium proceduralnym. Ciężar szczegółowej modalnej eksplikacji zależności przyczynowej, np. w kategoriach relacji (ukonieczniania) między uniwersaliami a partykulariami,⁶ składam – ku niezadowoleniu tego ostatniego – na barki zwolennika koncepcji koniecznościowej.⁷

⁶ Takie kroki poczynił Armstrong (1993; por. Psillos 2002, 167-72).

⁷ Podkreślmy, że niewywiązanie się z tego zadania w żaden sposób nie przekreśla roli eksplanacyjnej uogólnień projekcyjnych. Byłby jedynie fiaskiem realizacji tego zadania na

Armstrong w kolejnych uzupełnieniach wyjściowej koncepcji zaprezentowanej w (1983) wydaje się potwierdzać słuszność takiego programu i podaje nawet zarys rozwiązania problemu inferencji:

wyduje mi się oczywiste, że my (i zwierzęta) regularnie postrzegamy to, że jedna rzecz jest przyczyną innej. Osobiście wierzę, że przynajmniej niektóre z tych spostrzeżeń są tak pierwotne epistemicznie jak inne spostrzeżenia. Dobrym przykładem jest spostrzeżenie nacisku na nasze ciało.

Twierdzę zatem, że możemy rozpocząć od jednostkowych przyczynowań, które ujawniają pewien wzorzec, wzorzec, w którym – z grubsza rzecz biorąc – ten sam rodzaj przyczyn wywołuje ten sam rodzaj skutku. W tym właśnie miejscu można dokonać wnioskowania. Możemy wywnioskować istnienie uniwersaliów, aby wyjaśnić tę „samość”. (Armstrong 1993, 421)⁸

Zwolniony z obowiązku szczegółowego opracowania takiego programu oraz eksplikacji intensji dla zależności przyczynowej $X \rightarrow Y$, przechodzę do określenia jej ekstensji.

To ostatnie zadanie jest zasadniczo realizowane poprzez wprowadzone wcześniej kryterium proceduralne. Przypomnę, że zależność między zjawiskami uznamy za przyczynową na podstawie faktów niemodalnych, które określa odpowiedni wzorzec zależności statystycznych dla danych interwencyjnych tej zależności.

Kryterium proceduralne oraz odwołująca się do niego definicja uogólnień projekcyjnych pozwala w obrębie dyskutowanej tu koncepcji koniecznościowej wyodrębnić tę część rozwiązania zaproponowanego przez Armstronga, dla której problem inferencji jest rozwiązany. Zasadniczym bowiem celem kryterium proceduralnego jest uniemożliwienie stanowczych orzeczeń o zachodzeniu zależności przyczynowych w obszarze, gdzie nie uzyskano odpowiednich danych. W konsekwencji projektowana zależność przyczynowa $X : \rightarrow Y$ z definicji implikuje zachodzenie odpowiedniego wzorca zależności między zjawiskami X a Y (zgodnego z wymogami zależności przy-

gruncie koncepcji koniecznościowej. Stąd też możliwe jest takie przesunięcie *onus probandi* na zwolennika tej koncepcji.

⁸ Krytyczną dyskusję proponowanego tu przez Armstronga rozwiązania podaje (van Fraassen 1993). Strategia argumentacyjna w tym przypadku jest podobna jak w przypadku relacji ukonieczniania. Van Fraassen wyróżnia rodzaje przyczynowania osobno dla partykulariów i osobno dla uniwersaliów, analogicznie jak dla relacji ukonieczniania, pokazując brak zachodzenia związku logicznego w przypadku ogólnym.

czynowej $X \rightarrow Y$ wedle kryterium proceduralnego) tylko w tej dziedzinie, dla której pozyskano dane interwencyjne.⁹

W przeciwnym wypadku nie można przyjmować, że projekcja $X : \rightarrow Y$ automatycznie prowadzi do ustalenia zależności przyczynowej, gdyż byłoby to podważenie kryterium proceduralnego.

Podkreślimy więc, że przyjmując proponowane ograniczenie do klasy uogólnień projekcyjnych na gruncie koncepcji koniecznościowej, można otrzymać niekołowe i niezredukowane do pojęć niemodalnych określenie ekstensji uogólnień projekcyjnych i tym samym otrzymać rozwiązanie problemu inferencji w wyodrębnionym przez podejście proceduralne obszarze (dziedziny z danymi interwencyjnymi).

Pozostaje jednak niepewność co do wyników prac zwolennika koncepcji koniecznościowej nad intensją uogólnień projekcyjnych. A co jeśli ta wyprowadzi go poza określoną bezpiecznie ekstensję?

Wówczas otwarta jest jedna z dwóch możliwości. Jeśli wykazane zostaloby, że wyeksplikowanie intensji uogólnień projekcyjnych na gruncie koncepcji koniecznościowej nieuchronnie prowadzi do wyjścia poza wyznaczoną niezależnie ich ekstensję, wówczas należałoby oczekiwać, że takiej eksplikacji trzeba poszukiwać na gruncie innej ze zgłoszonych koncepcji praw przyrody.

Jeśli natomiast nie wykazano by tak mocnego związku, wówczas pozostaje przypomnienie wcześniejszego argumentu. W obszarze, który pozostaje poza ustaloną na gruncie proceduralnym ekstensją uogólnień projekcyjnych, między określonym wzorcem zależności pomiędzy badanymi zjawiskami a proponowanym uogólnieniem projekcyjnym nie ma związku ani logicznego, ani faktualnego. Zgodnie bowiem z przyjętą tu definicją uogólnienia projekcyjnego z projektowanej zależności $X : \rightarrow Y$ nie wynika logicznie zależność przyczynowa $X \rightarrow Y$ dla dziedziny, dla której nie posiadamy danych interwencyjnych. Z kolei posiadanie takich danych dla danej dziedziny gwarantuje zachodzenie zależności przyczynowej tylko pod warunkiem, że spełnione zostało w niej kryterium proceduralne.

Pozostawiam zwolennika koncepcji koniecznościowej z wyznaczeniem pewnych granic, jakie musi spełniać podane przez niego wytłumaczenie

⁹ Podobne rozwiązanie sugerował P. Feyerabend, podkreślając wzrost roli proceduralności w rozstrzygnięciu między modelami równoważnymi obserwacyjnie w obrębie danego wzorca (w sensie BN; por. Punkt 4.4) dla wzrastającej ogólności wiedzy naukowej (1962/1979, 104).

związku między zależnością przyczynową a kategoriami modalnymi, jak uniwersalia czy możliwości.

7.3 Uogólnienia projekcyjne a manipulacyjna koncepcja praw

Koncepcja manipulacyjna praw wyrasta z polemiki z koncepcją koniecznościową, a w szczególności z jej wersją zaproponowaną przez M. Tooleya (1997).¹⁰ Kategorię związku koniecznego między uniwersaliami oraz uniwersaliami egzemplifikowanymi w partykulariach Woodward (1992) zastępuje niezmienniczością. Przejawia się ona podczas interwencji, jakie są podejmowane przez człowieka lub samą przyrodę. Ugruntowaniem metafizycznym tej niezmienniczości mają być możliwość i moc poszczególnych systemów.

W rozwiązywaniu problemu identyfikacji, a więc odróżnianiu praw od przypadkowych uogólnień, istotną rolę odgrywają, zdaniem Woodwarda, wnioskowania kontrfaktyczne:¹¹

Twierdzić, że prawo grawitacji jest prawem przyrody, to twierdzić, że nie tylko poprawnie opisuje siły posiadane przez faktyczne masy, ale że mówi nam też coś na temat tego, jakie byłyby te siły w warunkach innych niż te, które faktycznie zachodzą. (Woodward 1992, 181-82)

Przyjętą przez siebie strategię realistyczną charakteryzuje jako zasadniczo różną od regularnościowej charakterystyki praw w odniesieniu do eksplikacji zawartości kontrfaktycznej praw. Zdaniem Woodwarda, zwolennicy koncepcji regularnościowej uznają element kontrfaktyczny za ugruntowany w zaobserwowanych regularnościach albo konwencjonalny. Realista natomiast tę treść modalną uznaje za „ugruntowaną w faktach świata przyrody” (1992, 182). Ten właśnie rodzaj ugruntowania różni prawa od przypadkowych uogólnień, pozwalając na potwierdzanie przez nie wnioskowań kontrfaktycznych.

Zasadniczym elementem krytyki koncepcji koniecznościowej przez Woodwarda – zwłaszcza w wersji zaproponowanej przez Tooleya – jest to, że nie „dostarcza żadnego redukcyjnego *wyjaśnienia* w kategoriach nieno-

¹⁰ Przykładem tradycyjnej charakterystyki praw przyczynowych, w której w ogóle nie uwzględnia się aspektu manipulacyjności, są poglądy Braithwaite'a (1955, 308-11).

¹¹ Bardziej dyskusyjne jest uznanie kontrfaktyczności za kryterium równoważne z uniwersalnością praw; por. (Krajewski 1998, 36-37) oraz uwagi we Wstępie o warunku kontrfaktycznym, który wprowadził Hume.

mologicznych, dlaczego prawa [...] potwierdzają zdania kontrfaktyczne czy opisują zależności, które są stabilne czy niezmiennicze w stosunku do zmian, jakie zachodzą w warunkach początkowych” (1992, 191). Zdaniem Woodwarda, ten brak ujawnia się szczególnie w zestawieniu zależności między uniwersaliami a partykulariami. Nie ma, jego zdaniem, w koncepcji koniecznościowej żadnego powodu, by przechodząc od jednego poziomu do drugiego, dokonywać jednocześnie przeskoku od zależności koniecznych do zależności przygodnych. To, że zależności konieczne obowiązują tylko między uniwersaliami, a nie partykulariami, Woodward uznaje arbitralnie przyjęte założenie, które należałoby dalej wyjaśnić. Z tej krytyki wyciąga on wnioski, że „analiza pojęć nomologicznych w kategoriach bardziej pierwotnych i lepiej rozumianych kategorii nie prowadzi do żadnego postępu” (1992, 191). Dlatego też porzuca zadanie redukcyjnej analizy praw.

Innym problemem koncepcji koniecznościowej, na jaki zwraca on uwagę, jest to, że prawa obowiązują tylko w pewnych dziedzinach, poza którymi niekiedy całkowicie się załamują. Nawet jednak w dziedzinach, w których obowiązują, nie są całkowicie precyzyjnym opisem systemu, którego dotyczą. Zawierają idealizacje, przybliżenia i uproszczenia. Tego nie można reprezentować po prostu jako zdania warunkowe, określające warunki obowiązywalności praw, gdyż są one często nie w pełni znane lub określone, są zbyt skomplikowane oraz byłoby to niezgodne z praktyką naukową formułowania praw. Nie rozwiązywałoby to też przybliżonej obowiązywalności praw w dziedzinach, w których obowiązują (Woodward 1992, 192).

W miejsce koncepcji koniecznościowej Woodward wprowadza więc własną, zgodnie z którą „w faktycznej praktyce naukowej pojęcie prawa jest blisko związane z pojęciami stabilności i niezmienniczości”. Prawo więc różni się od przypadkowego uogólnienia tym, że „wyraża zależność, jaka nie tylko zachodzi w faktycznych okolicznościach, ale która pozostanie stabilna czy niezmiennicza w stosunku do stosunkowo szerokiego zakresu zmian czy interwencji” (1992, 202). Pojęcie interwencji Woodward charakteryzuje ogólnie jako pojęcie kontrfaktyczne:

nie chodzi o to, czy my lub przyroda kiedykolwiek faktycznie doprowadzi do takich zmian [...], ale o to, co by się zdarzyło, gdyby (niezależnie od tego, jak) takie zmiany czy okoliczności miały się zdarzyć. (1992, 202)

Z uwagi na to, że pojęcie interwencji jest tu pojęciem kontrfaktycznym, Woodward podkreśla, że jego analiza praw przyrody nie jest redukcyjna.

Ograniczeniem, jakie Woodward wprowadza do powyższego ogólnego określenia praw, jest to, że prawa pozostają niezienne tylko w stosunku do określonej klasy interwencji – takie ograniczenie, jego zdaniem, pozostaje zgodne z tym, że „prawa nie muszą być i zwykle nie są zupełnie bezwyjątkowymi czy powszechnie prawdziwymi uogólnieniami” (1992, 202).

Proponowany przez Woodwarda związek ma także być zgodny z tym, że prawa przyrody – w odróżnieniu od przypadkowych uogólnień – są podstawą manipulacji zjawiskami przyrody i ich kontroli. Szczególnie jednak ten związek ujawnia się i jest doniosły w odniesieniu do zjawisk życia społecznego.

Metafizyczną podstawą praw rozumianych jako zależności niezmiennicze nie mogą być konieczne związki między uniwersaliami. Te zależności bowiem mają ograniczony zakres obowiązywalności, poza którym mogą się załamywać. To ujawnia – zdaniem Woodwarda – że podstawą praw są możliwości. Przysługują one przedmiotom bezwarunkowo, jednak mogą się manifestować tylko w określonych okolicznościach. To miałyby tłumaczyć, dlaczego charakterystyką praw jest niezmienniczość ograniczona do pewnej klasy interwencji. Takich interwencji, które nie prowadzą do zmian w okolicznościach uniemożliwiających ujawnianie możliwości przez przedmioty opisywane przez prawo.

Kluczowe dla zarysowanej tu koncepcji praw kategorii interwencji oraz niezmienniczości stały się przedmiotem dalszych prac Woodwarda (szczególnie 1997; 2000; 2003). Podając eksplikację pojęcia interwencji Woodward posiłkuje się kategorią przyczynowości, którą – jego zdaniem – adekwatnie reprezentują modele graficzne BN. Proponuje następującą definicję interwencji jako wyidealizowanego eksperymentu:

Zmienna I , przyjmująca określoną wartość $I = z_i$, jest interwencją na X względem Y wtedy i tylko wtedy, gdy I jest zmienną interwencyjną dla X względem Y oraz $I = z_i$ jest faktyczną przyczyną wartości przyjętej przez X . (2003, 98)

Zmienna I jest zmienną interwencyjną dla X względem Y , jeśli spełnia następujące warunki: 1) I jest przyczyną X ; 2) pewne wartości, które przyjmuje I , są takie, że kiedy I je przyjmuje, X przestaje zależeć od wartości innych zmiennych, które są przyczyną X , a zamiast tego zależy tylko od wartości przyjętej przez I ; 3) każda skierowana ścieżka od I do Y przechodzi przez X (I nie jest bezpośrednią przyczyną Y ani przyczyną bezpośrednich przyczyn Y poza X); 4) I jest statystycznie niezależne od dowolnej zmiennej Z , która jest przyczyną Y i leży na skierowanej ścieżce, która nie przechodzi przez X .

Trudno podać intuicyjny sens przyjętych przez Woodwarda określeń bez zwrócenia uwagi na pewną dwuznaczność, jaka tkwi u podstaw jego rozważań. Woodward zmierza nie tylko do określenia kryteriów pozwalających na odróżnienie praw od przypadkowych uogólnień, lecz jednocześnie chce skonstruować koncepcję, którą określa jako realistyczną. Realizacja tych dwóch zadań równocześnie prowadzi do wspomnianej wieloznaczności. W tej pracy, proponując podejście proceduralne, oddzielał te dwa zadania, które wprowadzają wieloznaczność w pracach Woodwarda. A mianowicie, za osobne zadanie uznaje podanie kryteriów zależności przyczynowych, które realizuję w tej pracy niezależnie od pojęciowej eksplikacji kategorii przyczynowości (Menzies i Price 1993, 196). Połączenie tych zadań przez Woodwarda skutkuje trudnościami w realizacji każdego z nich.

Tutaj zwróć uwagę na trudności w realizacji pierwszego zadania, jakim jest podanie kryteriów praw (a szczególnie zależności przyczynowych) w odróżnieniu od przypadkowych uogólnień. Nie wchodząc na razie w szczególności zaproponowanej przez Woodwarda definicji interwencji, zauważmy, że zakłada ona, iż określony już jest model zależności przyczynowych oraz znana jest definicja zależności przyczynowej lub jej kryteria. Przywołajmy więc odnośne definicje. Zacznę od pojęcia przyczyny całkowitej:

X jest całkowitą przyczyną *Y* wtedy i tylko wtedy, gdy istnieje możliwa interwencja na *X*, która zmieni *Y* lub rozkład prawdopodobieństwa *Y*. (2003, 51)

Zależność przyczynowa między *X* a *Y* jest zdefiniowana jako przyczyna bezpośrednia:

Warunkiem koniecznym i wystarczającym, aby *X* było (na poziomie typu) bezpośrednią przyczyną *Y* względem zbioru zmiennych *V*, jest to, aby istniała możliwa interwencja na *X*, która zmieni *Y* lub rozkład prawdopodobieństwa *Y*, kiedy pozostałe zmienne *Z_i* w *V* utrzymuje się zafiksowane na określonej wartości. (2003, 59)

Inny przypadek to taki, kiedy mowa jest o zależności między konkretną wartością zmiennej *X* a konkretną wartością zmiennej *Y* (na poziomie egzemplarza). Ten rodzaj przyczynowości Woodward określa jako przyczynowość faktyczną:

$X = x$ jest przyczyną *faktyczną* $Y = y$ wtw spełnione są dwa warunki: 1) zmienne *X* oraz *Y* faktycznie przyjmują wartości, odpowiednio *x* oraz *y*; 2) istnieje przynajmniej jedna ścieżka od *X* do *Y*, dla której interwencja na *X* zmieni wartość *Y*, pod warunkiem, że pozostałe bezpośrednie przyczyny *Z_i* zmiennej *Y* nieznajdujące się na tej ścieżce zostały zafiksowane na swoich faktycznych wartościach. (2003, 77)

Jeśli mowa o tym, że jedna zmienna jest przyczyną drugiej, to w grę wchodzi jedno z podanych wyżej określeń: jest ona przyczyną bezpośrednią lub pośrednią. Dokładniej, trzeba byłoby powiedzieć, że mówimy o zależności przyczynowej $X \rightarrow Y$ lub o wpływie X na Y w danym modelu przyczynowym i interesuje nas bezpośredni wpływ jednej zmiennej na drugą lub całkowity wpływ jednej zmiennej na drugą, gdzie mogą też pośredniczyć inne zmienne. Wbrew zatem sugestii Woodwarda, nie są to różne rodzaje przyczynowości, lecz *wpływu* czy efektu jednej zmiennej na drugą.

Uwagę zwraca występowanie w obu definicjach przyczynowości, zaproponowanych przez Woodwarda, kategorii interwencji – możliwej interwencji. W tego rodzaju analizie, jaką proponuje Woodward, trudno uniknąć kołowości i pozostaje tylko bronić zalet takiego podejścia, zwłaszcza że interwencja jest także pewnym rodzajem przyczynowości. Taką jednak obronę bardzo różnie ocenimy w zależności od tego, jakie zadanie ma ona zrealizować. Jeśli zadaniem, jakie chce tu rozwiązać Woodward, jest pojęciowa analiza praw (czy zależności przyczynowych), wówczas kołowość może być akceptowalna, pod warunkiem, że nie jest to koło bezpośrednie.

Inaczej jest natomiast w przypadku, gdyby realizowane zadanie miało doprecyzować kryteria odróżniania praw (zależności przyczynowych) od przypadkowych uogólnień. Otrzymujemy bowiem bezpośrednie koło w definiowaniu i intuicyjnie uchwytne sens zależności między przyczynowością a manipulacją znika w tych definicjach całkowicie.

Zamiast dyskutować na temat ewentualnych poprawek tych określeń, które lepiej oddałyby ten intuicyjny sens, zwróćmy uwagę na pewien element strategii Woodwarda, który wydaje się całkowicie odbierać sens takiemu przedsięwzięciu.

W definicji przyczynowości całkowitej oraz bezpośredniej wprowadza on kategorię *możliwej* interwencji. Eksplikując sens możliwości (2003, 128), Woodward podkreśla, że jest to możliwość fizyczna w tym sensie, że jej granice wyznaczone są przez prawa przyrody, ale są niezależne od warunków początkowych tych praw, jakie istnieją w naszym świecie. Taki manewr w realizacji zadania określenia kryteriów zależności przyczynowych wydaje się całkowicie niedopuszczalny. Uznanie bowiem, czy dany model zależności między zjawiskami jest przyczynowy, byłoby determinowane przez *analizę konceptualną* nierealizowalnych możliwości. A zatem uznanie, np. czy bakteria cholery była przyczyną epidemii w Londynie w XIX wieku, byłoby uzależnione nie, jak żądano, od uzyskania danych egzemplifikujących pewne obserwowalne regularności, ale od analizy tego, co byłoby, gdy prze-

prowadzić takie a takie manipulacje w wyidealizowanych warunkach eksperymentalnych.

To pokazuje także zasadniczą słabość w przyjętym przez Woodwarda pojęciu interwencji. Z jednej strony bowiem, cytowana powyżej definicja interwencji – pomijając wprowadzoną w niej kołowość – wydaje się nakładać na tę kategorię niesprawdzalne warunki. Określenie, czy zmienna X przyjęła określoną wartość z uwagi na przyjęcie określonej wartości przez inną zmienną jest niemożliwe bez przyjęcia określonego modelu przyczynowego. Z kolei określenie, czy dany model jest przyczynowy, wymaga zastosowania pewnych narzędzi analizy statystycznej – niezależnie od wymogów nakładanych przez podejście BN, RCM czy proponowane tu proceduralne – do zebranych danych dla zjawisk reprezentowanych przez zmienne losowe X oraz Y .

Z drugiej zaś strony, Woodward poszerza pojęcie interwencji w perspektywie możliwości, gdzie staje się ona w zasadzie niezwiązana z działaniem innym niż *analiza konceptualna*. Takie osłabione pojęcie interwencji z kolei dopuszcza daleko posuniętą dowolność w interpretowaniu tego, co można uznać za możliwą interwencję.

Inny jeszcze sposób wskazania głębokiego pęknięcia w podstawach koncepcji Woodwarda podsuwa wypowiedź M. Hellera. Jak zaznaczyłem wcześniej, Woodward uznaje interwencję za możliwą w tym sensie, że jest ona zgodna z prawami, ale niezależna od ich warunków początkowych. W sedno tego stwierdzenia trafia – zapewne niezamierzony – komentarz:

Zależności przyczynowe należy tu rozumieć w sensie jak najbardziej fizycznym; przyczyną jest na przykład wysłanie promienia świetlnego, a skutkiem – zarejestrowanie tego promienia na kliszy lub w fotokomorcie. Tego rodzaju zależności przyczynowe *zawsze można przetłumaczyć na warunki początkowe* pewnych równań i na ogół tak nie jest, że warunki początkowe determinują całą przeszłość i przyszłość. Wszystko zależy od geometrii czasoprzestrzeni. W tym znaczeniu przyczynowość staje się aspektem geometrii. (M. Heller 2002, 27; podkreślenie – P. K.)

Przypomnę, że możliwa interwencja jest podstawą dla Woodwarda w definiowaniu innych rodzajów interwencji oraz różnych rodzajów zależności przyczynowych. Po wyeliminowaniu warunków początkowych, zgodnie z sugestią Hellera, Woodwardowi pozostaje wyłącznie ten rodzaj analizy zależności przyczynowych, który w przyпыływie sceptycyzmu przyczynowego zawładnął umysłem B. Russella (1912/1913). Chodzi mianowicie o zaproponowaną przez niego konceptualną analizę przyczynowości, redukującą zależności przyczynowe do funkcyjnych.

Zwróćmy też uwagę, że druga z kluczowych dla koncepcji Woodwarda kategorii, a mianowicie niezmienniczość, jest definiowana względem – pewnej klasy – interwencji, co pogłębia jeszcze problem. To, które z obserwowanych lub stwierdzonych zależności uznane zostaną za niezmiennicze, wydaje się podwójnie relatywne. Po pierwsze, jak pokazywałem wyżej, nie ma jasnych kryteriów wyodrębnienia interwencji. W związku z tym pozostajemy wyłącznie przy intuicyjnym i czysto subiektywnym pojęciu o bardzo chwiejnym znaczeniu.

Po drugie, określenie klasy interwencji, z uwagi na którą ma być zachowana niezmienniczość, też nie jest określone w koncepcji Woodwarda precyzyjnie. Jak zwrócił uwagę Psillos (2002, 183-84), każde przypadkowe uogólnienie jest także niezmiennicze względem pewnej klasy interwencji.

Wobec powyższego powstaje istotny problem. Woodward proponuje zastąpienie praw przyrody przez klasę uogólnień niezmienniczych. Bowiem za warunek posiadania mocy eksplanacyjnej uznaje on niezmienniczość uogólnień. Z powyższych uwag wynika jednak, że tę klasę różni od uogólnień przypadkowych jedynie różnica *stopnia*. Stajemy więc wobec arbitralności, której – jak pokazałem w poprzednim punkcie – nie usunie metafizyczna kategoria możliwości.

7.4 Uogólnienia projekcyjne a relatywizm w kwestii praw

Znacznie większą rolę niż w dotychczas dyskutowanych koncepcjach ukryte metafizyczne własności odgrywają w N. Cartwright koncepcji względności praw przyrody. Proponuje ona redukcijną analizę tych praw, które uznaje nie za ściśle ogólne, lecz dopuszczające wyjątki. Posługując się kategorią wprowadzoną w modelowaniu ekonometrycznym przez A. Marshalla, o wszystkich prawach twierdzi, że są *ceteris paribus*, a więc że obowiązują w pewnych tylko okolicznościach, o ile pozostałe warunki nie podlegają zmianie. I tak, np. ciała zachowują swój stan ruchu – zgodnie z pierwszym prawem Newtona – jeśli nie oddziałuje na nie siła zewnętrzna.¹²

Koncepcja Cartwright modyfikuje problem identyfikacji. Regularności, które nie dopuszczałyby wyjątków, nie istnieją. Zatem prawa przyrody nie

¹² Zupełnie przeciwnie charakteryzuje ten przykład Heller: „Do pojęcia przyczynowości fizyka Newtonowska wprowadziła istotne modyfikacje. Ruch jednostajny (inercyjny) przestał być procesem, który wymagałby przyczyny; stał się stanem, który może trwać dowolnie długo, i jedynie zmiana stanu (przyspieszenie) wymaga przyczyny (siły)” (2002, 25).

mogą być szczególnym rodzajem takich regularności. Jedyny rodzaj regularności, jaki obserwujemy, to regularności względne w stosunku do określonych (niezmieniających się) okoliczności. Jeśli zatem obserwujemy w świecie zachodzenie regularności, to ich podstawą zawsze są byty metafizyczne określane przez Cartwright jako maszyny nomologiczne. Charakteryzuje je ona jako:

(wystarczająco) zafiksowany układ składników lub czynników o (wystarczająco) stabilnych możliwościach, które w odpowiedniego rodzaju (wystarczająco) stałym środowisku, przy powtarzalnym działaniu, będą generować taki rodzaj regularnego zachowania, które reprezentujemy w prawach naukowych. (Cartwright 1999, 50)

Elementem maszyn nomologicznych, który ostatecznie jest odpowiedzialny za generowanie regularności w obserwowanym zachowaniu są możliwości. One decydują o tym, „co przedmioty mogą zrobić” (Cartwright 1995, 156). Możliwości manifestują się w postaci regularności tylko jednak w przypadku, gdy są elementem maszyn nomologicznych.

Obecność maszyn nomologicznych sprawia więc, że prawdziwa jest formuła *ceteris paribus*. Eliminują bowiem działanie tych czynników, które zakłóciłyby obserwowaną regularność. W tym sensie więc wszystkie prawa przyrody są względne: „są prawdziwe tylko *ceteris paribus* – tylko względem skutecznie powtarzanego działania maszyny nomologicznej” (1999, 49-50).

System planetarny może być przykładem takiej maszyny nomologicznej, która spontanicznie występuje w przyrodzie. Cartwright podkreśla jednak, że w większości przypadków takie maszyny są dziełem człowieka, powstającym w sztucznych warunkach laboratoryjnych.

Zastanawiając się nad spójnością stanowiska Cartwright, Psillos (2002, 193-95) dochodzi do wniosku, że kluczowe dla tej koncepcji pojęcie możliwości prowadzi do rozbieżnych wniosków. Z jednej strony bowiem, możliwości mają być mierzalne, a twierdzenia naukowe zawierające przypisywanie możliwości przedmiotom, testowalne. Z drugiej jednak strony, możliwości nie redukują się do regularności, czego najważniejszym dla Cartwright przejawem jest ich ujawnianie się w pojedynczych przypadkach. W tym drugim przypadku, jak podkreśla Psillos, trudno jednak stosować ideę mierzalności możliwości oraz testowalności, gdyż – wzięwszy pod uwagę, że możliwości do ujawnienia wymagają także zaistnienia odpowiednich okoliczności – można rzekome ujawnienie możliwości w jednostkowym przypadku potraktować jako zjawisko losowe.

Sądzę, że zasadnicza różnica między statusem praw przyrody a przypisywaniem możliwości przedmiotom tkwi w kwestii względności. Przedmioty posiadają daną możliwość lub nie w sposób bezwzględny. Jedynie manifestacja tej możliwości jest uzależniona od okoliczności, w jakich ulokowany jest ten przedmiot: może być to manifestacja jednorazowa bądź regularna, gdy możliwość występuje jako element maszyny nomologicznej.

Prawa przyrody natomiast dla Cartwright są zawsze względne. Obowiązują zawsze pod klauzulą *ceteris paribus*, która jest zachowana o tyle, o ile istnieje maszyna nomologiczna.

Zasadnicze pytanie dotyczy tego, czy przypisanie możliwości w istotny sposób modyfikowałoby wyjaśnienie przez podanie uogólnienia projekcyjnego? Podobny problem w swojej analizie koncepcji Cartwright podjęła M. Morrison (1995). Bez określenia specjalnej procedury, która pozwalałaby przypisywać możliwość przedmiotowi na podstawie jednokrotnej manifestacji, jesteśmy zdani na znane skądinąd procedury określania zależności przyczynowych między poszczególnymi zjawiskami. Jeśli przyjąć, jak chciałaby Cartwright, że prowadzą one tylko do wniosków obowiązujących *ceteris paribus*, wówczas status przypisywanych możliwości nie mógłby być inny: przypisanie możliwości przedmiotowi także byłoby względne.

Jeszcze bardziej problematyczne wydaje się dopuszczenie możliwości przypisywania przedmiotom możliwości, które w ogóle nie zostały ujawnione. W takim przypadku, jak słusznie zauważa Psillos (2002, 197), trudno jest pogodzić ten pogląd z tezą Cartwright, że możliwości są mierzalne, a zdania naukowe przypisujące możliwości przedmiotom – testowalne.

Koncepcja względności praw przyrody podnosi pewien kluczowy problem dla proponowanego tu stanowiska proceduralnego. Na problem ten zwrócono uwagę w dyskusji na tematy istnienia praw bardziej podstawowych niż względne (Earman i in. 1999, 2002) oraz redukowalności praw różnych poziomów (van Fraassen, 1993a).

W swoim komentarzu do zarysu koncepcji Cartwright van Fraassen (1993b) krótko charakteryzuje dwie możliwe strategie odpowiedzi na względność praw przyczynowych. Jedna strategia, którą szczegółowo realizują Earman i in. (1999; 2002), dopuszcza względność praw w obrębie grupy praw wtórnych czy wyprowadzonych. Obowiązują one względnie wedle reguły *ceteris paribus*, ale są również wyprowadzone z praw fundamentalnych. Względność nie jest charakterystyką istotną praw, lecz wiąże się jedynie z ich wtórnością wobec praw fundamentalnych. Tę strategię określeń jako redukcijną.

Druga strategia, którą wyróżnia van Fraassen, to strategia dystrybucyjna. W zarysowanej przez niego za R. Collingwoodem wersji restryktywnej tej strategii pokazuje się, że prawa przyczynowe nie są względne, jednak obowiązują tylko dla pewnego wybranego poziomu dyskursu. Pozostałe natomiast poziomy dyskursu traktuje się jako niezwiązane z tym wyróżnionym.¹³

Przyjęta przez Earmana i J. Robertsa (1999) strategia redukcyjna zakorzeniona jest w koncepcji praw, którą daje najlepszy system. Należy oczekiwać, że nauka u kresu dedukcyjnie uporządkuje wszystkie prawa w taki sposób, że wyróżniony zostanie pewien rdzeń, który stanowić będą, zdaniem Earmana i Robertsa, fundamentalne prawa fizyki teoretycznej. Argumentują oni, że takie prawa obowiązują bezwzględnie, a więc bez klauzuli *ceteris paribus*. Pozostałe prawa, a więc fenomenologiczne prawa fizyki (dotyczące przedmiotów makro) oraz nauk szczegółowych, są wtórne i jako takie dopuszczają również występowanie klauzuli *ceteris paribus*.

W pełnej rozciągłości strategia redukcyjna Earmana i Robertsa nie może być bezpośrednio odniesiona do uogólnień projekcyjnych. Jednym z powodów są dyskutowane w Punkcie 7.1 trudności regularnościowej koncepcji praw. Zasadniczy jednak powód dotyczy zależności zachodzących między poszczególnymi uogólnieniami projekcyjnymi. Zgodnie z przyjętym ich rozumieniem, są one względne w stosunku do przyjętego modelu przyczynowego. Relacje pomiędzy poszczególnymi modelami nie mają charakteru związku logicznego, czego wymagałoby „porządkowanie” narzucone przez strategię Earmana i Robertsa.

W przypadku modeli przyczynowych porządkowanie dokonuje się poprzez agregację zmiennych. Okazuje się jednak, że tylko w bardzo szczególnych warunkach taka operacja zachowuje zależności przyczynowe stwierdzone dla agregowanych zmiennych (Chu i in. 2003).

Dla modelu zatem, który powstaje jako agregacja zmiennych wcześniej znanych modeli przyczynowych, należy osobno stosować kryterium proceduralne, niezależnie od spełniania go dla zmiennych w modelach „niższego” stopnia hierarchii. Przy ustalaniu zatem, kiedy dana zależność między zmiennymi jest przyczynowa oraz kiedy obowiązuje dane uogólnienie projekcyjne, nie można zastosować strategii redukcyjnej, proponowanej przez Earmana i Robertsa. W przyjętym przez nich sensie nie ma „fundamentalnych” zależności przyczynowych, z których można byłoby – po przyjęciu dodatkowych założeń dotyczących danej dziedziny – wyprowadzić zależno-

¹³ Por. Zakończenie. Wersję nierestryktywną na gruncie probabilistycznej koncepcji przyczynowości rozwijają systematycznie (Price i Menzies 1993).

ści przyczynowe „wtórne”. Uogólnienie projekcyjne sformułowane na podstawie przyjętego modelu obowiązuje więc niezależnie od związków logicznych, jakie zachodzą między zmiennymi w różnych modelach przyczynowych.

Strategia redukcyjna Earmana i Roberta wydaje się jednak naświetlać pewien istotny problem związany ze strategią dystrybucyjną, jaką do eksplikacji intensji przyczynowości stosują H. Price i P. Menzies (1993). Earman i Roberts, podobnie jak Cartwright, sugerują, że w przypadku wszystkich praw przyczynowych mamy do czynienia z względnością, ujawnianą przez występowanie klauzuli *ceteris paribus*. A zatem ta właśnie warunkowość, jeśli eksplikujemy „przyczynowości” dla różnego rodzaju obserwowanych regularności, stanowi nieodłączny element takiej eksplikacji.

Dodatkowym powodem jest to, że modele przyczynowe muszą zachowywać specyfikację probabilistyczną w postaci błędu losowego. W przeciwnym wypadku pogwałcone zostałyby założenie wierności (por. Punkt 4.4). Pomijając względy aplikacyjne, opuszczenie w specyfikacji modelu elementu błędu losowego prowadziłoby do podania modelu przyczynowego w postaci deterministycznej. Z tym jednak wiązałyby się wzajemna przekładalność funkcjonalna poszczególnych zmiennych, a w konsekwencji uniemożliwienie z *zasady* jednoznacznego stwierdzenia zachodzenia relacji przyczynowej na podstawie danych. Z założenia bowiem dopuszczone byłyby wzajemnie równoważne reprezentacje, spośród których nie można byłoby wskazać jednej.

Przyjęcie, że klauzula *ceteris paribus* jest nieodłącznym elementem eksplikacji intensji zależności przyczynowych, prowadzi jednak do nieintuicyjnych wniosków. Pierwszy z nich dotyczy niezgodności z charakterem podstawowego epistemicznie doświadczenia – „percepcji” – przyczynowości, na którą zwracają uwagę Cartwright i Armstrong. Taka percepcja, która ich zdaniem należy do klasy najbardziej elementarnych, nie wydaje się nieuchronnie zawierać żadnego elementu charakterystyki warunkowej czy losowej.¹⁴

Innym powodem nieintuicyjności jest tu aprioryczne wykluczenie pewnych możliwych eksplikacji intensji przyczynowości. Zasadniczą grupą takich eksplikacji wykluczonych przez przyjęcie klauzuli *ceteris paribus* jako nieodłącznej charakterystyki zależności przyczynowych, jest grupa eksplika-

¹⁴ Intensjonalność pojęcia przyczynowości dyskutują szerzej (Mackie 1974; von Wright 1974, 6-7; Anscombe 1971/1981; Bobryk 1992; Żegleń 2003, 174-204).

cji określanych wspólnie zbiorczym mianem teorii sprawczych.¹⁵ Najpełniej rozwiniętą postać takiej teorii zaprezentowali Price i Menzies (1993),¹⁶ przyjmując pierwotność charakterystyki zależności przyczynowej w odniesieniu do ludzkiej sprawczości, a następnie reinterpreterując kategorię przyczynowości stosowaną w innych kontekstach jako własność wtórną, projektowaną na podstawie doświadczenia sprawczości. Nie przesądzam tu, że taka teoria jest właściwą eksplikacją intensji zależności przyczynowej, jednak wprowadzenie klauzuli *ceteris paribus* wydaje się arbitralnie zamykać drzwi takiej możliwości jeszcze przed zbadaniem jej wewnętrznej spójności.

Trzecim powodem, na który chciałbym zwrócić tu uwagę, jest взгляд historyczny. Przywołajmy dyskutowaną szerzej w części pierwszej rozbieżność między reprezentowanym przez Snowa deterministycznym podejściem do przyczynowości a losowym, które forsowane było przez członków the Board of Health. Snow rozumiał zależność przyczynową między bakterią cholery a chorobą deterministycznie: jeśli ta bakteria została wprowadzona do organizmu, wywołuje ona chorobę. Przypadki, kiedy tak nie jest, należy wytłumaczyć, odwołując się do okoliczności zaburzających normalny proces rozwoju choroby. Jakkolwiek deterministyczne rozumowanie Snowa nie jest sformułowane jawnie, to żądanie wprowadzenia klauzuli *ceteris paribus*, wydaje się prowadzić do nieadekwatności historycznej. Ponadto, ta nieadekwatność historyczna ma dalsze konsekwencje w postaci zaburzenia rekonstrukcji różnic stanowisk Snowa i the Board of Health. Narzuca ona przyjętą przez tę instytucję kategoryzację analizowanej zależności przyczynowej.

Ponieważ zadaniem tej pracy nie jest podanie eksplikacji intensji zależności przyczynowej, ograniczę się więc do uwagi, że klauzula *ceteris paribus* nie może być elementem wprowadzonym do charakterystyki każdej zależności przyczynowej. Earman i Roberts wydają się podsuwać rozwiązanie tego problemu. Można stwierdzić analogicznie, że przy eksplikacji intensji zależności przyczynowej da się wyodrębnić pierwotną i wtórną grupę zależności z uwagi na pewną wiążącą je relację R, szczegółowo ilustruje to

¹⁵ T. Czeżowski (1933, 8-9) charakteryzuje takie stanowisko jako fenomenalistyczne (przyczyną jest fakt, proces lub zjawisko) oraz idealistyczne (związek przyczynowy jest „wyrazem sposobu, w jaki umysł nasz ujmuje przebieg zjawisk”). Por. też (Krajewski 1967, 35-51).

¹⁶ G. von Wright, który zainicjował tego typu stanowisko, nie bronił go w całej rozciągłości, uznając walentność dychotomii: przyczynowość (przyrodnicza) – sprawczość (podmiotowa); por. zvl. (von Wright 1974, 49-54).

podejście w (Price i Menzies 1993). Wprowadzenie charakterystyki podstawowej grupy zależności z pominięciem klauzuli *ceteris paribus* nadal pozostaje zgodne z tym, że relacja R, jaka łączy pozostałe zależności przyczynowe z grupą podstawową, prowadzi do modyfikacji charakterystyki tej grupy wtórnej o klauzulę *ceteris paribus*.

Względność praw przyrody więc, którą w swoich pracach wyakcentowała Cartwright, nie jest powszechnym udziałem uogólnień projekcyjnych, a jedynie po spełnieniu omówionych wyżej warunków. Z kolei lokalność związana z określaniem ich obowiązywalności w danej dziedzinie nie jest w żaden sposób uzależniona od przyjęcia postulowanych przez nią kategorii metafizycznych, jak możliwości czy maszyny nomologiczne.

7.5 Uogólnienia projekcyjne a sceptycyzm w kwestii praw

W skrajnej postaci stanowisko sceptycyzmu co do istnienia praw przyrody zaprezentował van Fraassen w (1989). Podejmuje on polemikę z regularnościową i koniecznościową koncepcją praw przyrody. Zasadniczą rolę w tej krytyce odgrywa „duet” problemów identyfikacji i inferencji. Wiele z opartych na nich zarzutów przedstawiłem w poprzednich punktach, omawiając te dwie koncepcje praw.

Przypisywanie istnienia prawom przyrody van Fraassen zasadniczo traktuje jak reifikację języka, który ma odniesienie wyłącznie do sposobu konstrukcji modeli naukowych. Podstawową rolę odgrywa tu zasada symetrii: problemy symetryczne pod istotnym względem powinny mieć takie same rozwiązania. Grupę przekształceń zachowujących symetrię problemu charakteryzuje się poprzez podanie stałych dla tych przekształceń. Prawa więc, o których mowa w nauce, określają sposób konstruowania modeli naukowych, a poszukiwanie dla nich odniesienia przedmiotowego jest niczym innym, jak reifikacją wyrażen językowych. Miejsce praw przyrody w dotychczasowych analizach nauki zajmują, określane przez van Fraassena nie-reifikująco, podstawowe zasady nauki, które są zależnościami współzmiennymi dla wszystkich układów odniesienia danej grupy przekształceń (van Fraassen 1989, 285). Pod względem epistemicznym prawa są akceptowane (przyjmowane z innych względów niż prawdziwość, zwłaszcza informatywność), natomiast przekonanie co do prawdziwości dotyczy modeli naukowych o tyle, o ile odnoszą się do tego, co obserwowalne (adekwatność empiryczna). Przekonanie to nie dotyczy natomiast jakiejś „ukrytej rzeczywistości”, którą miałyby reprezentować reifikowane prawa.

Po tym zarysie sceptycznej koncepcji van Fraassena przejdę do jej szczegółowego omówienia, przede wszystkim koncepcji zależności współzmiennych, by następnie wyświecić różnicę między nimi a proponowanymi w podejściu proceduralnym uogólnieniami projekcyjnymi.

Rozpocznę od przykładu, który podaje van Fraassen (1989, 279). Niech F' oznacza układ poruszający się z prędkością v względem osi X układu F w kierunku dodatnim. Współrzędna X położenia ciała zmienia się zgodnie z zależnością s : $x' = x - vt$. Jak zmienia się prędkość tego ciała? Załóżmy, że w ciągu $t = 2$ w F osiąga ono punkt o współrzędnych (a, b) . Zgodnie z przekształceniem s w F' to ciało osiąga punkt o współrzędnych $(a - 2v, b)$. Średnia droga przebyta przez to ciało w F wynosi więc $a/2$, natomiast w F' wynosi $(a - 2v)/2 = a/2 - v \neq s(a/2) = a/2 - 2v$. Wniosek van Fraassen wyciąga więc następujący: „Oczywiście prędkość nie jest niezmiennicza i przekształcana jest *systematycznie* wraz z położeniem, ale *nie w ten sam sposób*” (1989, 279).

Za H. Weylem autor *Laws and symmetry* wprowadza ideę współzmienności: dana wielkość może zmieniać się wraz ze zmianą współrzędnych dla niektórych ich przekształceń. Niech G będzie grupą wszystkich liniowych przekształceń normalnych systemów współrzędnych w przestrzeni lub czasoprzestrzeni. Wielkość Q o n składowych $Q_F = (a_1, \dots, a_n)$ względem dowolnego systemu współrzędnych F jest określana jako współzmienna względem G , jeśli $Q_{F'} = sQ_F$ dla F' otrzymanego z F przez dowolne przekształcenie s z G .

W tym sensie prędkość, w omawianym wcześniej przykładzie, jest współzmienna względem grupy obrotów, ale nie względem całej grupy transformacji Galileusza.

Pojęcie współzmienności van Fraassen z zastosowania do wielkości odnosi do „teorii, praw, zasad, twierdzeń i równań”. Nazywając równoważnymi wszystkie układy odniesienia otrzymane przez przekształcenia z danej grupy, definiuje współzmiennność dla równań następująco: „Równanie $Q_F(x) = f(Q^1_F(x), \dots, Q^n_F(x))$ jest *współzmiennie* dokładnie wtedy, gdy jego wartość logiczna jest taka sama we wszystkich równoważnych układach odniesienia” (van Fraassen 1989, 280).

Ogólną postać pojęcia współzmienności otrzymujemy w przypadku, gdy dowolny układ odniesienia można otrzymać z pozostałych przez zastosowanie przekształceń z danej grupy: „Równanie jest *współzmiennie* dokładnie wtedy, gdy jest albo prawdziwe we wszystkich układach odniesień albo w żadnym” (van Fraassen 1989, 281).

Z dwóch zatem wyróżnionych przez niego jako powszechnie uznawane cechy praw, a mianowicie ogólności i konieczności, van Fraassen uznaje ogólność za adekwatnie reprezentowaną przez współzmiennność: „Współzmiennność zatem jako zastosowana do równania jest odmianą *ogólności*. Jest to ten rodzaj ogólności, którego oczekujemy od rzekomych praw przyrody – albo, rezygnując z metafory, od wszystkich podstawowych zasad nauki” (van Fraassen 1989, 281).

Współzmiennie równania nie są jednak dokładnym zastępnikiem rzekomych praw przyrody. Pozostawiając „zbyt wiele niezmienniczych”, dopuszczalne przez współzmiennność przekształcenia nie pozostawiają miejsca na tradycyjnie przypisywaną prawom konieczność. Ta klasa przekształceń bowiem jest zbyt bogata, by jednoznacznie wyodrębnić jeden spośród możliwych światów oraz zagwarantować, że zmiany zgodne z nimi – w przeciwieństwie do praw przyrody – nie wyprowadzą nas poza świat aktualny (van Fraassen 1989, 288-89).

Współzmiennność, jak twierdzi van Fraassen, nie jest trywialną czy powierzchowną charakterystyką ogólności. Uogólnienie „Wszystkie rudziki są czerwone” lub „sód zawsze spala się na żółto” mają syntaktyczną postać zdań ogólnych. Jednak z grona uogólnień dyskwalifikuje je, zdaniem van Fraassena, kryterium współzmienności. To bowiem, co jest czerwone dla obserwatora w jednym układzie odniesienia, może mieć inny kolor dla obserwatora w innym układzie odniesienia (poruszającym się z odpowiednią prędkością względem pierwszego).

Jednak to, co uznaje się za klasę dopuszczalnych przekształceń, jest elementem uzależnionym od teorii. A zatem pewne twierdzenia, które uznano by za ogólne w fizyce klasycznej, w fizyce relatywistycznej nie mogą mieć tego statusu.

Po dyskusji na temat pojęcia twierdzeń współzmiennych jako „zastępnika” prawa przyrody przejdźmy z kolei do pytań o ich związek z wprowadzonymi w tej pracy uogólnieniami projekcyjnymi. Czy te ostatnie są ogólne w sensie van Fraassena? I czy powinny być? Przy odpowiedzi na to pytanie powrócę jeszcze do innych fragmentów analiz van Fraassena.

Zaproponowana przez van Fraassena eksplikacja pojęcia ogólności praw czy zasad nauki poprzez ekstrapolację kategorii współzmienności intuicyjnie może wydawać się trafna. Od ogólności żądamy bezwyjątkowej obowiązywalności, co w terminologii van Fraassena przekłada się na obowiązywalność równania we wszystkich równoważnych układach odniesienia. Jak podkreśla van Fraassen, w jego ujęciu ta bezwyjątkowa obowiązywalność

jest jednak *wyłącznie* kwestią przyjętej w nauce reprezentacji zjawisk, a nie np. przejawem jakiejś „głębszej” niż obserwowana tylko rzeczywistości, jaką miały być prawa przyrody.

Czy współmienność nie darzy nas jednak nadmiarem symetrii? Dlaczego bowiem nie miałyby być tak, że ogólność w sensie van Fraassena jest wykrojona tak, aby „uratować” ogólność wszystkich syntaktycznie ogólnych zdań, np. przykładowego zdania o rudzikach.

Natychmiastowe rozwiązanie, jakie się tutaj nasuwa, to odpowiednie skonstruowanie grupy przekształceń, względem której oceniana jest równoważność układów odniesień. Wybranie trywialnej grupy, jaką jest grupa identyczności, przekształcająca układy w te same układy, gwarantuje nam, że wszystkie syntaktyczne uogólnienia są współmienne, a tym samym ogólne w sensie van Fraassena. Różnica między uogólnieniami gramatycznymi a „głębszymi” znika...

Taką propozycję można by jednak uznać za banalizację rozwiązania oferowanego przez van Fraassena. Grupa przekształceń, a tym samym klasa dopuszczalnych układów odniesienia musi być szersza niż tylko identyczność, o czym decydują względy teoretyczne. Oto jak van Fraassen we własnych słowach charakteryzuje tę różnicę:

Prawdziwa ogólność nie może leżeć w formie słów użytych do wyrażenia sądu – a zatem na czym ona polega? Ujmę to następująco: czy prawem może być *wszystkie rudziki są czerwone* lub *sód zawsze spala się na żółto*? W teorii fizycznej poszukiwany rodzaj ogólności to *niezmienniczość* dla wszystkich symetrii. A zatem odpowiedź musi brzmieć: nie, nie może być prawem. Ponieważ kolor nie jest niezmienny dla transformacji Galileusza, podobnie jak prędkość. To, że wszystkie rudziki są czerwone tak samo nie może być prawem, jak to, że każde ciało niepoddane działaniu siły jest w spoczynku. (van Fraassen 1989, 275)

Dalsze wyjaśnienie ma pokazać, dlaczego zmienia się kolor rudzika wraz z przejściem do innego układu odniesienia. Kolor mierzy się jako częstotliwość światła odbijanego lub emitowanego przez dany obiekt. Światło sferycznie przenosi się jako fala. Fala o długości λ zależy od częstotliwości f następująco: $\lambda = v/f$. W układzie F' poruszającym się względem wyjściowego z prędkością k otrzymujemy inną częstotliwość: $f' = f(v - k)/v < f$. Ten efekt znany jest jako przesunięcie ku czerwieni. Wniosek, jaki wyciąga van Fraassen w tym miejscu, jest następujący:

Oczywiście fizyczna konstytucja światła w żaden sposób nie jest zmieniona przez obserwowanie go przez poruszającego się obserwatora. Przesunięcie

jest efektem perspektywicznym – kolor nie jest niezmienniczy, jest jedynie względny. (van Fraassen 276)

Co z tego wynika dla rudzików? Czy obserwator z poruszającego się układu F' może powiedzieć, że rudziki są fioletowe? Wyobraźmy sobie sytuację, gdzie jedyny układ odniesienia zawierający rudziki jest układem, który: 1) nie zawiera żadnego obserwatora oraz 2) porusza się z prędkością k względem pozostałych układów odniesienia. Zdanie „Wszystkie rudziki są czerwone” jest prawdziwe tylko względnie, a mianowicie w postaci „Wszystkie rudziki są czerwone względem obserwatora w układzie F' ”.

Nawet jeśli dać odpowiedź pozytywną na to pytanie, to nadal pozostaje kwestia tego, w jaki sposób dobrać klasę układów odniesienia dla układu, w którym przebywają rudziki. Nie jest jasne, dlaczego taka klasa miałaby być wyznaczona przez – dyskutowaną we wszystkich przykładach współzmienności w *Laws and symmetry* klasa transformacji Galileusza. Jakie względy teoretyczne powstrzymałyby przed tym, aby uznać, że układy poruszające się z określoną prędkością w stosunku do układu odniesienia z rudzikami nie są dopuszczalne? Ogólniej, dlaczego nie można poza klasą układów dopuszczalnych umieścić wszystkich, które powstają jako efekt stosowania przekształceń prowadzących do zmiany obserwowanego koloru?

Brak na to pytanie wyraźnej odpowiedzi w *Laws and symmetry*. Spróbuję jednak ją zrekonstruować. Uogólnienia, które są proponowane na poszczególnych etapach rozwoju nauki, są współzmiennie względem pewnej klasy równoważnych układów odniesienia. Nie posiadają jednak tej własności względem innych układów. Przykładem są prawa mechaniki Newtona, które nie są współzmiennie względem układów odniesienia fizyki relatywistycznej. Przejście od fizyki klasycznej do relatywistycznej pokazuje jednak, że idealny kierunek rozwoju nauki polega na poszerzaniu klasy równoważnych układów odniesienia, względem których dane (w postaci zmodyfikowanej) uogólnienie jest prawdziwe. Jeśli zatem jakieś syntaktyczne uogólnienie jest prawdziwe względem ograniczonej klasy układów odniesienia oraz z zasady nie może być prawdziwe względem poszerzonej klasy, należy je wykluczyć z grona uogólnień „prawdziwych”.

Jak pokazuje argumentacja Earmana (1978, 177-79), takie jednak rozumowanie może być słuszne tylko w odniesieniu do absolutnej czasoprzestrzeni, jako wyróżnionego układu odniesienia. Jeśli przyjąć, że ostatecznie klasa równoważnych układów odniesienia ma objąć wszystkie możliwe, to należy zapytać, co oznacza to ostatnie. Jeśli przyjąć, że klasa wszystkich możliwych układów odniesienia to klasa logicznie możliwych układów od-

niesienia, wydaje się, że warunek współzmienności spełniałyby wyłącznie tautologie i ostatecznie otrzymalibyśmy kolejną trywializację proponowaną przez van Fraassena pojęcia ogólności.

Inna propozycja brzmiałaby: jest to klasa *fizycznie* możliwych układów odniesienia. Taka jednak klasa byłaby determinowana przez uogólnienia w sensie van Fraassena, a więc przez uogólnienia współzmiennie. Te zaś są definiowane przez zachowywanie wartości logicznej w układach fizycznie możliwych. Nie można bronić się w tym miejscu, jak podkreśla Earman, twierdząc, że ta kołowatość jest niewinna, gdyż na początku wystarczy znać wyłącznie te uogólnienia współzmiennie, które determinują strukturę czasoprzestrzeni (układów odniesienia):

w teoriach takich, jak ogólna teoria względności, gdzie struktura czasoprzestrzeni staje się elementem dynamicznym, prawa, które determinują strukturę czasoprzestrzeni zwykle są powiązane ze wszystkimi pozostałymi prawami. (Earman 1978, 177-78)

Trudno zatem, jak podkreśla Earman, wyeksplikować ogólność jako cechę *pojedynczych* zdań. Jej miejsce raczej zajmie charakterystyka pewnego zespołu powiązanych twierdzeń. W takim jednak przypadku ogólność nie wydaje się warunkiem koniecznym praw, którego nie zastąpiłyby inne, jak zrównoważenie prostoty i informatywności (Earman 1978, 179).

Zatem ruzdzikom nie można zarzucać, że z *zasady* są czerwone tylko w wybranych układach. Każde bowiem uogólnienie współzmiennie jest także prawdziwe tylko w wybranych układach. O tym, jaki powinien być ten dobór, decydują względy teoretyczne, ale trudno – jak pokazałem za Earmanem – oczekiwać, że van Fraassen podałby niekołowatą charakterystykę tej determinacji teoretycznej.

Być może da się jednak obronić kategorię współzmienności, dodając, że teoretyczne zdeterminowanie klasy dopuszczalnych układów odniesienia musi dodatkowo być odniesione do innego kryterium, jakim jest adekwatność empiryczna. Te zatem zasady teoretyczne, które determinują klasę dopuszczalnych układów odniesienia muszą być adekwatne empirycznie. Zatem nie dowolne uogólnienie syntaktyczne jest uogólnieniem w sensie van Fraassena, gdyż wykrojenie klasy dopuszczalnych układów odniesienia przez zasady teoretyczne jest zasadniczo ograniczone adekwatnością empiryczną tych ostatnich.

Taka próba obrony oznaczałaby rewizję kategorię stwierdzenia, że ogólność jest wyłącznie własnością reprezentacji. Nawet przy zgodzie van Fraassena na taką rewizję, szanse na powodzenie tej koncepcji wydają się

jednak nikle. Problem Goodmana jest dokładnie zaadresowany do przypadku, kiedy zachowujemy adekwatność empiryczną, a jednocześnie reprezentacja dopuszcza *nadmiar* symetrii. Problem Goodmana wymaga niejako odwrócenia porządku dotychczasowej dyskusji. Dotąd zastanawialiśmy się nad tym, czy kryterium współzmienności wyodrębnia z klasy uogólnień syntaktycznych takie, które są uogólnieniami „prawdziwymi”, uogólnieniami w sensie głębszym. Problem Goodmana natomiast prowadzi do wniosku, że dla dowolnego uogólnienia w sensie głębszym – przy zachowaniu adekwatności empirycznej – jesteśmy w stanie podać jego zmodyfikowaną wersję, która jest wyłącznie uogólnieniem syntaktycznym. I na odwrót, dla każdego uogólnienia syntaktycznego można podać taką jego postać, która spełnia warunek współzmienności.

Ostatecznie więc wniosek, do jakiego prowadzi problem Goodmana, jest taki, że do zdefiniowania ogólności w sensie głębszym nie wystarczy poszerzenie pojęcia współzmienności jako własności reprezentacji o kryterium adekwatności empirycznej. Zakorzenie w praktyce językowej nie wydaje się rozwiązaniem akceptowalnym dla van Fraassena, podobnie jak nie jest nim zrównoważenie prostoty i informacyjności, które poddał krytyce w (1989).

W dotychczasowej dyskusji zakładałem, że znany jest sposób wyodrębniania przekształceń, którym mogą być poddawane układy odniesienia, a także te ostatnie. Van Fraassen oferuje dość mglisty sposób postępowania:

W argumentacji na podstawie symetrii postępuje się następująco: stwierdza się problem, a następnie podejmuje pierwszą próbę wyizolowania zasadniczych cech tej sytuacji problemowej, a więc cech, które są istotne dla rozwiązania. Na ogół nie jest konieczny głęboki wgląd: dokładne stwierdzenie problemu na ogół określa to, co zamierzone. [...] Wyizolowanie istotnej czy relewantnej struktury jest równoważne zdefiniowaniu zbioru przekształceń, które pozostawiają ten problem zasadniczo taki sam. Te przekształcenia to *symetrie* tego problemu. [...] [G]dy istotne parametry są wyizolowane, rozwiązanie musi polegać na podaniu reguły (tzn. funkcji), która zależy tylko od tych parametrów. (1989, 257-58)

W przypadku rudzików wiadomą rzeczą jest, że wprowadzenie przekształcenia polegającego na zmianie prędkości pociąga za sobą także zmianę ich koloru. Można, jak się wydaje robić to van Fraassen, przyjąć mocną tezę o redukcjonizmie własności występujących w dowolnym dyskursie do własności fizycznych. Wówczas nie ma powodu, by wśród przekształceń nie znalazło się takie, gdyż – jak podkreśla van Fraassen – nie zmienia ono „fizycznej konstytucji” światła, względem której kolor jest wtórny.

Nie negując fizycznej wtórności światła, można – przynajmniej w zasadzie – bronić autonomiczności dyskursu biologicznego czy zoologicznego, w którym sformułowane zostało uogólnienie „Wszystkie rudziki są czerwone”. Funkcjonalna rola koloru czerwonego w przypadku rudzików w żaden sposób nie jest bezpośrednio wyprowadzalna z konstytucji fizycznej tego koloru. Nie jest więc jasny powód, dla którego należałoby przyjmować – jak *implicite* zdaje się to robić van Fraassen – tezę o pierwotności dyskursu fizycznego.

W odpowiedzi na komentarze w dyskusji wokół *Laws and symmetry* van Fraassen zarysował inną jeszcze strategię. Zgodnie z nią przyjęcie mocnej tezy o redukowalności wydaje się zbędne. Mówiąc bowiem o uogólnieniach prawdziwych, dokonujemy rekonstrukcji uogólnień w nauce teoretycznej. W tym przypadku wszystkie modele obejmują wyłącznie własności punktów czasoprzestrzeni. Natomiast sposób identyfikacji rudzików oraz ich koloru na podstawie modeli nauki teoretycznej to jest zagadnienie nauki stosowanej oraz koniecznych założeń.

Problem redukcji wystąpi jednak ponownie, chociaż być może w innej płaszczyźnie. Nie tylko bowiem stwierdzenia nauki teoretycznej, o której pisze van Fraassen, będą posiadały ogólność, lecz także nauki stosowanej. Przyjmując, że podana w kategoriach współzmienności eksplikacja dotyczy w pierwszym rzędzie stwierdzeń nauki teoretycznej, natrafiamy na problem redukcji. W jaki sposób, bez przyjęcia dodatkowych założeń, pokazać, że ogólność wszystkich twierdzeń nauk stosowanych wywodzi się z tej wyeksplikowanej pierwotnie dla nauki teoretycznej? Wydaje się, że jest to niemożliwe bez przyjęcia założenia, które określałoby, że symetrie problemu sformułowanego w nauce teoretycznej nie prowadzą do naruszenia „odpowiednika” tego problemu w nauce stosowanej. Mocna teza, że takie symetrie są zachowane we wskazanej przez van Fraassena nauce teoretycznej w przypadku dowolnego uogólnienia dowolnej nauki stosowanej, wydaje się jednak odpowiednikiem wcześniej dyskutowanej mocnej tezy redukcyjnej.

Zwróćmy uwagę na inny jeszcze problem, jaki można zauważyć, odnosząc powyżej przytoczony cytat do przywoływanych w tej pracy fragmentów dziejów nauki. Wyodrębnienie istotnych parametrów, względem których określane są symetrie i niezmienniczość problemu, nie jest, jak sugerowałby van Fraassen w przytoczonym tekście, tylko kwestią językową czy konceptualną. W rzeczywistości jest to bowiem złożony i długotrwały proces, w którym istotną rolę odgrywa także interakcja badawcza z rzeczywistością. Badając przyczyny epidemii cholery w Londynie w latach 1850.,

działająca z ramienia parlamentu komisja the Board of Health przyjęła koncepcję miazmatyczną. Wprowadzała ona – jako część charakterystyki problemu – symetrię między cholera a obserwowanymi objawami pacjentów, np. bakteriami, które około 30 lat później Koch zidentyfikował jako przyczynę cholery. Taka symetria pozostawała w niezgodzie z wyciągniętymi przez Snowa wnioskami na podstawie analizy symptomów i rozwoju cholery u pacjentów. Przyjęta przez van Fraassena strategia charakterystyki ogólności na podstawie symetrii narzuconych przez badany problem wydaje się nie ujmować pewnej charakterystyki ogólności, która jest istotna dla ich roli eksplanacyjnej. Nawet dodatkowe uwzględnienie kryterium adekwatności empirycznej, lub mocniejszego jeszcze – skuteczności podejmowanych interwencji – nie wydaje się w tym przypadku poprawiać perspektyw analizy podejmowanej przez van Fraassena. Zarówno bowiem komisja, jak i Snow, posiadali adekwatne empirycznie uogólnienia, które w obu przypadkach były podstawą skutecznego zapobiegania nowym przypadkom zachorowań. Przyjęta przez van Fraassena strategia eksplikowania substytutu praw przyrody w kategoriach współzmienności pomija co najmniej jeden czynnik, który okazuje się zasadniczy przy określaniu roli eksplanacyjnej tych podstawowych zasad nauki. Tym elementem, który uwzględniają wprowadzone w tej pracy uogólnienia projekcyjne, jest uwzględnienie odpowiedniego rodzaju danych przy formułowaniu uogólnienia. Takie dane w przypadku modeli przyczynowych, jak pokazywałem wcześniej, muszą pochodzić z pola interwencji danej zależności i mieć charakter danych interwencyjnych. Bez ich uwzględnienia niemożliwe wydaje się podanie zadowalającej eksplikacji uogólnień, która zachowywałaby związek z ich rolą eksplanacyjną. Fakt, że to Snow podał wyjaśnienie epidemii cholery, a nie komisja, wydaje się poza zasięgiem proponowanych przez van Fraassena uogólnień współzmiennych.

Zastanawiając się zatem nad proponowanym przez van Fraassena sposobem determinowania klasy „dopuszczalnych przekształceń”, dochodzimy do dwóch wniosków. Po pierwsze, grupa transformacji Galileusza, do której odnoszą się wszystkie ilustracje w *Laws and symmetry*, oraz ogólne zalecenia są ograniczone do jednego rodzaju dyskursu, jakim jest dyskurs fizyczny. Bez przyjęcia jakiejś wersji mocnej tezy o redukowalności innych postaci dyskursu do tego wyróżnionego, nie można w żaden bezpośredni sposób przenieść kategorii współzmienności na pozostałe uogólnienia.

Po drugie, nawet przyjęcie mocnej tezy o redukowalności nie gwarantuje, że kategoria współzmienności jest adekwatną eksplikacją prawdziwej – a nie

czysto syntaktycznej – ogólności. W przypadku rudzików bowiem nadal można bronić stanowiska, że dopuszczenie przekształcenia prędkości układów względnych jest taką klasą przekształceń, która nie zachowuje „symetrii problemu”. Kluczowy bowiem predykat w podanym uogólnieniu to „jest czerwony”, a takie przekształcenie prowadzi do zmiany symetrii problemu. W układach, które poruszają się względem danego, obserwowany będzie efekt przesunięcia ku czerwieni, a zatem zmieniona będzie symetria problemu *przy takim* przekształceniu. Samo więc sformułowanie problemu wydaje się nasuwać wniosek, że takie przekształcenie należy wykluczyć z klasy dopuszczalnych przekształceń właśnie *na podstawie* mocnej tezy redukcyjnej.

Samo uogólnienie kategorii współzmienności nasuwa także podejrzenia co do jej stosowalności w kontekście zaproponowanym przez van Fraassena. Przypomnę, że Weyl odnosił ją do wielkości fizycznych i zmian, którym podlega przy poszczególnych przekształceniach. Jeśli przez zastosowanie danego przekształcenia bezpośrednio do danej wielkości można było otrzymać tę wielkość w takiej postaci, w jakiej występowała ona w układzie odniesienia, który podlegał temu przekształceniu, to Weyl uznawał ją za współzmienną.

Odniesienie tej kategorii do równań wydaje się jednak budzić pewne wątpliwości. Załóżmy, że dana funkcja jest zdefiniowana jako monotonicznie malejąca w przedziale wartości (x_0, x_i) oraz jako monotonicznie rosnąca w przedziale wartości (x_j, x_n) , gdzie $x_i < x_j$. Załóżmy też, że istnieje takie przekształcenie s , że $s: (x_0, x_i) \rightarrow (x_j, x_n)$. Ponadto przyjmę, że w wyjściowym układzie odniesienia nie jest możliwe zrealizowanie wartości spoza przedziału (x_0, x_i) . Nazwanie takiego równania ogólnym wydaje się niezgodne z naszym intuicyjnym rozumieniem uogólnienia, podobnie jak w przykładzie ziebieskich trudno mówić o uogólnieniu indukcyjnym. Podany przykład jest bardziej jeszcze problematyczny, gdyż rzeczą – z założenia – wiadomą jest niemożność zrealizowania w danym układzie przedziału funkcji, w którym jest ona monotonicznie rosnąca. Mimo tego, zgodnie z podaną przez van Fraassena definicją, jest to równanie współzienne, a więc ogólne w sensie prawdziwym.

Powyższą listę trudności dla zaproponowanego przez van Fraassena substytutu praw naukowych uzupełnę o skonstruowany przez niego samego duet problemów identyfikacji i inferencji. W tym przypadku w problemie identyfikacji pytamy o to, jak wyodrębnić klasę współzmiennych i prawdziwych uogólnień od uogólnień czysto syntaktycznych. W problemie inferen-

cji – jak z uogólnień współzmiennych wyprowadzić wszystkie zaobserwowane regularności.

Wcześniej poczynione uwagi nasuwają wniosek, że van Fraassen nie podaje zadowalającego odróżnienia uogólnień syntaktycznych od prawdziwych. Takie odróżnienie podał wyłącznie dla kontekstu fizycznego, ale nawet tutaj jest ono, jak pokazują przytoczone przeze mnie oraz sformułowane przez Earmana argumenty, problematyczne.

Przywołany tu wniosek z Earmana analizy kategorii ogólności praw przyrody ujawnia też problem inferencji w rozwiązaniu van Fraassena. Nie można bowiem, jak się wydaje, podać charakterystyki ogólności praw, nie odnosząc się do jakiejś ich zbiorczej charakterystyki takiej, jaką na gruncie koncepcji regularnościowej jest zrównoważenie prostoty i informatywności. Eksplicacja zatem ogólności jako własności poszczególnych stwierdzeń jest uwikłana w pośrednią kołowość i z tego też powodu nie pozwala na pokazanie, w jaki sposób wynikają z niej zaobserwowane regularności.

Podsumuję ten punkt, odnosząc uogólnienia projekcyjne do zaproponowanej przez van Fraassena kategorii współzmienności. Jak pokazałem w Rozdziale 6, uogólnienia projekcyjne zawierają wystarczającą charakterystykę ogólności dla określenia roli eksplanacyjnej.

Uogólnienia projekcyjne mają w zapleczu model przyczynowy, a więc są pochodną interakcji z badaną rzeczywistością. Jako takie nie są, w odróżnieniu od uogólnień współzmiennych, wyłącznie pochodnymi przyjętego modelu. Dzięki temu wyodrębniają tę klasę uogólnień, która pozostaje w istotnym związku z wyjaśnianiem, mimo że nie spełnia mocnych warunków, jakie zwykle nakłada się na prawa przyrody. W tym też sensie uogólnienia projekcyjne zachowują podstawowe intuicje, które wyznaczają przyjęte przez van Fraassena rozwiązanie. Nawet gdyby kategoria współzmienności dostarczała zadowalającej eksplicacji ogólności, nie miałoby to wpływu na rolę eksplanacyjną, jaką pełnią uogólnienia projekcyjne. Otrzymalibyśmy być może walentną filozoficznie charakterystykę uogólnień, nieprzybliżającą nas jednak do określenia ich funkcji eksplanacyjnej. Niezależnie zatem od listy trudności, jakie wskazałem dla propozycji van Fraassena, ten projekt wydaje się ukierunkowany w inną stronę niż ta, w którą zmierzają podjęte tu rozważania.

Zależności przyczynowe w wyjaśnianiu subsumpcyjnym

Zagadnienie wyjaśniania ma w filozofii długą historię. Wyartykułowane koncepcje filozoficzne zaczęły pojawiać się dopiero w XX wieku. Mimo dzielących je różnic i rozbieżności wspólnym punktem odniesienia jest sięgająca starożytności idea, że wyjaśnić jakiś fakt to znaczy pokazać jego związek z pewną ogólną prawidłowością.¹ Ten rodzaj koncepcji wyjaśniania będę określał tu jako *subsumpcyjny*.²

Mniej zauważalnym punktem odniesienia współczesnych dyskusji wokół zagadnienia wyjaśniania jest inna, także szczytująca się historią sięgającą

¹ Kamiński w systematyce rozumowań wskazał trzy podstawowe rodzaje wyjaśnień: przez uniwersalizację, teoretyczne (określane tu łącznie jako subsumpcyjne) oraz wyjaśnianie przez porównanie zdarzenia humanistycznego z typowym (1981/1994, 245), często eksplikowane w kategoriach rozumienia. Ten trzeci z rodzajów wyjaśniania zostaje poza prowadzoną tu dyskusją. Próby te zwięźle charakteryzują (Bronk 1988, 99-120; Majdański 1972, 164-75; Marciszewski 1963). Pominięta w tej pracy jest także dyskusja na temat różnic między wyjaśnianiem przyczynowym deterministycznym a teleologicznym (Metallmann 1929/2002; 1934; Bocheński 1954/1993; von Wright 1971; Krajewski 1976, 61-80). Posługując się jednak za J. Życińskim (2001, 297) kategorią ukierunkowania strukturalnego, do tej dychotomii można odnieść wprowadzoną w Punkcie 9.3 koncepcję wyjaśniania jako subsumpcji proceduralnej.

² Wskazana tu subsumpcyjna koncepcja wyjaśniania może być potraktowana jako szczególny przypadek wyartykułowanej przez K. Ajdukiewicza koncepcji wyjaśniania jako wyrowadzenia danego zdania (eksplanandum) ze zdań wcześniej uznanych; por. (Ajdukiewicz 1964/1985, 405-407).

antyku, koncepcja. Zgodnie z nią wyjaśnić fakt to znaczy podać jego przyczyny. Ten rodzaj koncepcji wyjaśniania będą określał tu jako *przyczynowy*.³

Niewątpliwie model przyczynowy traktowano w XX wieku jako ukryte odniesienie dla wyjaśniania.⁴ Zasadniczym powodem tego był brak zaproponowanego we Wprowadzeniu odróżnienia filozoficznego od kryterialnego podejścia do zagadnienia przyczynowości⁵ oraz konsekwencje tego braku. Kategorię przyczynowości traktowano jako wyłącznie filozoficzną i niepodatną na wyeksplikowanie, zwłaszcza takie, które precyzyjnie określałoby kryteria zależności przyczynowych. Podkreślimy, że proponowane tu odróżnienie podejść filozoficznego od kryterialnego nie dewaluuje tego pierwszego. Przeciwnie, należałoby oczekiwać, że filozoficzna koncepcja zależności przyczynowych pozwoli pełniej uchwycić sens zależności przyczynowych oraz wyjaśniania.⁶

Analizując wybrane współczesne koncepcje wyjaśniania, zwrócę uwagę na związki koncepcji subsumpcyjnej oraz koncepcji przyczynowej wyjaśniania. Postaram się pokazać, że podstawowe trudności koncepcji wyjaśniania wynikają z niedostatecznego uwzględnienia koncepcji przyczynowej⁷ lub nieprzestrzegania odróżnienia podejścia filozoficznego od kryterialnego.⁸

³ W podobny sposób zasadniczy podział współczesnych koncepcji wyjaśniania wprowadza (Kałuszyńska 1995, 27).

⁴ Roli wyjaśnień przyczynowych w naukach humanistycznych (historii) bronili A. Malewski i J. Topolski (1960, 6-7 oraz 114-57). A ogólnie m.in. J. M. Bocheński: „[...] w wielu naukach [...] [p]rzeważa raczej wyjaśnianie kauzalne, które polega na podaniu przyczyny fenomenu” (1954/1993, 115).

⁵ Koresponduje ono z drugim i trzecim poziomem analiz w filozofii i metodologii nauk wyróżnionymi przez E. Kałuszyńską (1994, 183-87).

⁶ Obie te koncepcje występują w Arystotelesa koncepcji wyjaśniania (Hankinson 1998, 161-167; Losee 1993, 9-12), której mniej istotnym założeniem było, aby przesłanki były prawdziwe, niedowodliwe oraz lepiej znane od wniosku (*Analityki wtóre* 71 b 20 – 72 a 5).

⁷ Zdaniem Czeżowskiego (1933, 28-31), podporządkowanie koncepcji wyjaśniania przyczynowego wyjaśnianiu subsumpcyjnemu należałoby przypisać Platonowi, uznającemu idee za przyczyny (celowe) przedmiotów, dziedziczących jednocześnie związki logiczne zachodzące między realizowanymi ideami.

⁸ Z tego też powodu nietrafny wydaje mi się pogląd Amsterdamskiego, który uznał spór o koncepcję wyjaśniania (zwłaszcza w naukach społecznych) za metodologicznie i filozoficznie nierozstrzygalny, gdyż uzależniony od akceptacji określonego ideału wiedzy naukowej (1983, 134).

Osobno dyskusji poddam koncepcję wyjaśniania przyczynowego, jaką współcześnie zaproponował Woodward.⁹ Uwzględnia ona najnowsze wyniki prac nad graficznym modelowaniem zależności przyczynowych BN. Można by w związku z tym oczekiwać, że w zadowalający sposób respektuje także odróżnienie podejścia filozoficznego od kryterialnego. Pokażę jednak, że trudności zauważone przy analizie pojęcia interwencji Woodwarda, mają poważne konsekwencje dla zaproponowanej przez niego koncepcji wyjaśniania przyczynowego. Zwrócę uwagę, że w tej koncepcji zarówno zagadnienie proceduralności zależności przyczynowych, jak też projekcyjności uogólnień nie są wolne od trudności.¹⁰

W związku z tym zaproponowana zostanie koncepcja wyjaśniania przyczynowego wolna od tych trudności. Pokażę też, że wyjaśnianie przyczynowe jest istotnym elementem metodologii proceduralnej, gdyż integruje proceduralność zależności przyczynowych z projekcyjnością uogólnień. Zasadnicza idea, która jest eksplikowana w ostatnim rozdziale, może być zarysowana następująco. Wyjaśnienie danego faktu, zgodnie z koncepcją subsumpcyjną, polega na wskazaniu jego związku z ogólną prawidłowością. Tę prawidłowość, zgodnie z koncepcją przyczynową, wyraża – spełniający kryterium proceduralne – model przyczynowy. Zgodnie z koncepcją subsumpcyjną otwarta jest jednak możliwość, że ta prawidłowość daje się uogólnić projekcyjnie, dzięki czemu otrzymalibyśmy pełniejsze wyjaśnienie. Określenie więc, że taka prawidłowość obowiązuje w nowej dziedzinie, wymaga wykazania, że w niej zaproponowany model przyczynowy spełnia kryterium proceduralne. Metodologia proceduralna uchwytuje w ten sposób swoistą dynamikę postępowania badawczego.¹¹

8.1 Zależności przyczynowe w wyjaśnianiu DN oraz IS

Za W. Salmonem, najwybitniejszym współczesnym teoretykiem wyjaśniania, powtórzę:

⁹ Na osobną natomiast okazję pozostawiam systematyczne rozważenie odniesienia współczesnych dyskusji do koncepcji wyjaśniania przyczynowego M. Webera, którą przypomnieli Kmita i Nowak (1968, 110-17), a której związek z proponowaną tu koncepcją wyjaśniania sygnalizuję krótko w Punkcie 9.3.

¹⁰ Bardziej systematyczny przegląd możliwych w takim przypadku trudności podał Przełęcki (1959/1966, 111-20).

¹¹ Por. też (Czeżowski 1968, 192).

Poszukiwanie naukowego zrozumienia świata ma korzenie sięgające czasów starożytnych znacznie poprzedzających Arystotelesa. Mimo że dalekie od założeń, to *naukowe* poszukiwanie okazało się bardzo owocne w wielu dziedzinach badań. Filozofowie i naukowcy, co najmniej od czasów Arystotelesa, starali się też określić, na czym takie zrozumienie polega. To *filozoficzne* poszukiwanie aż do całkiem niedawna było znacznie mniej udane. Twierdzą jednak, że w okresie ostatnich czterech dekad dokonał się znaczący postęp filozoficzny. (1990, 1)

Początek wyróżnionym przez Salmona okresom dał przełomowy artykuł C. G. Hempela i P. Oppenheima pt. „Studies in the logic of explanation” (1948). Późniejsze dyskusje i nowsze koncepcje wyjaśniania nawiązują do zaproponowanej w tym tekście koncepcji. Jej też poświęcę więcej uwagi.

W tej koncepcji¹² wyartykułowano wyjaśnianie subsumpcyjne w postaci koncepcji dedukcyjno-nomologicznej (DN). To, co ma być wyjaśnione, czyli eksplanandum, jest opisem (jednostkowego) zdarzenia. Eksplanans, czyli to, co wyjaśnia, zawiera zasadniczo dwa elementy: prawo¹³ (gr. *nomos*, stąd nomologiczny) oraz opis warunków początkowych, w których zaistniało wyjaśniane zdarzenie. Wyjaśnianie uznaje się za zadowalające przede wszystkim wtedy, gdy eksplanandum wynika logicznie z tak skonstruowanego eksplanansu (stąd dedukcyjny):

rozważane zdarzenie jest wyjaśnione przez subsumpcję pod ogólne prawa, czyli przez pokazanie, że zaszło zgodnie z tymi prawami dzięki realizacji pewnych określonych warunków początkowych. (Hempel i Oppenheim 1948, 136)

Oprócz tego żąda się, by podane prawo zawierało treść empiryczną (było confirmowalne przez doświadczenie), a eksplanans by był prawdziwy. Tylko ten ostatni warunek, jaki mają spełniać zadowalające wyjaśnienia, Hempel uznawał za „empiryczny”, czyli pozalogiczny oraz niesyntaktyczny. Wynikanie logiczne, prawo naukowe i confirmowalność traktował jako pojęcia logiczne (1948, 137-38), które można zdefiniować w oparciu o samą formę logiczną zdań, czyli czysto syntaktycznie. Jeśli wyjaśnienie spełnia tylko trzy pierwsze, ale nie czwarty warunek, to jest „wyjaśnieniem potencjalnym”, a nie „faktycznym”.

¹² Uwzględniam tu także późniejsze prace C. G. Hempela (zwl. 1965). Por. też (Krajewski 1998, 44-48).

¹³ Współcześnie sygnalizuje się tu słabszy wymóg, jakim jest zdanie prawo-podobne, a więc niekoniecznie spełniające wszystkie warunki nakładane na prawa naukowe, a w szczególności pod takie określenie podpadają także hipotezy naukowe.

Wzorując się na sukcesach logiki w klaryfikacji podstaw matematyki, autorzy dedukcyjno-nomologicznej koncepcji wyjaśniania uznali logikę za zasadnicze narzędzie w analizie pojęć naukowych (Psillos 2002, 220). Powstała w ten sposób – krótko opisana wyżej – charakterystyka wyjaśniania, która była w zasadzie strukturalno-syntaktyczna. Miało to ugruntować obiektywność wyjaśniania naukowego poprzez jednoznaczne ujawnienie jego logicznej formy oraz wykazanie, że może być sensownie aplikowana.¹⁴ Wyjaśnianie w tym ujęciu jest szczególnego rodzaju prawomocnym wnioskowaniem dedukcyjnym, a odróżnia je występowanie w przesłankach praw naukowych.¹⁵

Przyjęte przez Hempla rozwiązanie jest szczególną postacią koncepcji subsumpcyjnej, którą krótko można wyrazić następującą formułą:

wyjaśnić to wydedukować z prawa naukowego. (D)

W Hempla ujęciu koncepcji subsumpcyjnej szczególnym przypadkiem jest koncepcja przyczynowa wyjaśniania:

Wyjaśnianie przyczynowe jest szczególną odmianą wyjaśniania dedukcyjno-nomologicznego. O pewnym zdarzeniu lub ich zbiorze można powiedzieć, że jest przyczyną określonego „skutku” tylko wtedy, gdy istnieją prawa ogólne je łączące w taki sposób, że z opisu zdarzeń poprzedzających wystąpienie skutku, może być wydedukowane za pomocą praw. (1959/1965, 300-301)

Istnieją natomiast wyjaśnienia dedukcyjno-nomologiczne, które nie są przyczynowe.

Takie ujęcie wyjaśniania przyczynowego ma u podstaw specyficzne rozumienie zależności przyczynowej.¹⁶ To, że jedno zdarzenie jest przyczyną drugiego, należy rozumieć według Hempla jako stwierdzenie, że istnieją

¹⁴ E. Mokrzycki podkreślał normatywny charakter proponowanej w koncepcji DN logicznej rekonstrukcji struktury wyjaśniania, gdyż jej nieaplikowalność w stosunku do pewnych rodzajów nauk, np. socjologii, traktowano jako podstawę do ich dewaluacji; por. (Mokrzycki 1980, 159-60).

¹⁵ Wielu autorów podkreśla, że to pojęcie wyjaśniania jest przede wszystkim *epistemiczne* – wyjaśnić jakieś zdarzenie to wykazać, że należało się go spodziewać, wzięwszy pod uwagę prawa rządzące zachodzeniem tego rodzaju zdarzeń oraz warunki wstępne, w których ono faktycznie zaistniało.

¹⁶ Uwagi poczynione przez K. Poppera (1934/2002, 53-55) można odczytać jako tendencję przeciwstawną w tym względzie koncepcji Hempla i bliższą rozważanej w tej pracy. Por. też (Hajduk 1970, 95-96).

prawa, ze względu na które wystąpienie zdarzenia wcześniejszego jest nomologicznie wystarczające dla wystąpienia zdarzenia późniejszego.

W tym poglądzie Hempła łączą się dwa elementy. Jeden można określić jako determinizm przyczynowy, drugi jako totalny deskrypcjonizm. Załóżmy, że znamy wszystkie prawa, które odnoszą się do okoliczności, w których zaistniało wyjaśniane przez nas zdarzenie. Podanie tych praw oraz wyliczenie wszystkich odnośnych warunków początkowych gwarantuje, że skutek wynika logicznie. W tym sensie więc jest on zdeterminowany przez obowiązujące prawa.

Równie mocne wydaje się założenie Hempła co do tego, że jesteśmy w stanie wyliczyć wszystkie odnośne prawa oraz warunki początkowe w taki sposób, aby zagwarantować zachodzenie relacji dedukowania skutku z praw. Pominięcie którejkolwiek z przesłanek prowadziłoby bowiem do braku żądanej relacji logicznej.

W tym względzie postępowanie Hempła przypomina to, jak Carnap próbował podać reguły logiki indukcji, żądając totalnych dowodów oraz językowego ujęcia całości ludzkiej wiedzy, obecnej i przyszłej (Kawalec 2003a, 122-25). W obu przypadkach wydaje się, że jest to próba dopasowania rzeczywistości do wymogów nakładanych przez formę logiczną.

Niespełnienie mocnych założeń, jakie wprowadził Hempel, ma jednak ważne konsekwencje dla proponowanej przez niego koncepcji wyjaśniania. Jeśli z zasady nie jest możliwe sprostanie wymogom totalnego deskrypcjonizmu, lub – jeśli jest, ale nadal przyczynowość ma charakter losowy, niedeterministyczny, to koncepcja przyczynowa wyjaśniania *nie jest* szczególnym przypadkiem koncepcji subsumpcyjnej. Wyartykułowanie formy wyjaśniania subsumpcyjnego w postaci koncepcji DN okazało się, jak pokazują też przywołane niżej jej trudności, niewystarczające.

Podstawowe bowiem zarzuty przeciw tej koncepcji, kwestionujące jej adekwatność jako koncepcji wyjaśniania, można uznać za konsekwencję braku niezależnego opracowania formy wyjaśniania przyczynowego i próbę zredukowania jej – jako logicznej konsekwencji – do wyjaśniania subsumpcyjnego (Psillos 2002, 224; Topolski 1984, 465-83).¹⁷ Eksplícite taką redukcję na gruncie koncepcji DN dyskutował K. R. Popper (1962/1992, 96-97).

¹⁷ Zestawiając koncepcję Arystotelesa ze współczesnymi, Hankinson komentuje: „To, w jakim stopniu potrafimy określić właściwy kierunek wyjaśniania, będzie zależało od tego, czy uda nam się pokazać, że jedna własność rzeczywiście zależy przyczynowo od drugiej. A to z kolei można odczytać z tego, w jaki sposób nasza interwencja zmieniałaby daną sytuację. [...] Współczesne koncepcje, odrzucające pojęcie przyczynowości jako anachronizm, który nale-

Pierwsza grupa zarzutów, wykazujących, że koncepcja DN nie stanowi warunku wystarczającego wyjaśniania, może być zilustrowana przykładem formułowanym mniej więcej następująco. Zacznę od wyjaśnienia zgodnego z DN: wysokość pnia i kąt względem Słońca jako warunek wstępny oraz prawa optyki geometrycznej (wraz z elementarną trygonometrią) składają się na eksplanans, z którego wynika logicznie długość cienia. S. Bromberger (1966) zaproponował odwrócenie tego przykładu: przy długości cienia i określonym kącie względem Słońca jako warunku wstępnym oraz tych samych prawach można logicznie wydedukować wysokość pnia. Zgodnie z DN należałoby zatem uznać, że nie tylko długość pnia wyjaśnia długość jego cienia, ale też na odwrót. To ostatnie, choć zgodne z DN, razi swoją niezgodnością z intuicją, gdyż to nie cień powoduje, że pień ma taką lub inną wysokość.

Uogólniając, koncepcja DN dopuszcza symetryczność między zdaniem opisującym przyczynę oraz opisującym skutek – jedno może wyjaśniać drugie i na odwrót.¹⁸ Związek przyczynowy jest jednak asymetryczny i dlatego DN nieadekwatnie ujmuje wyjaśnianie przyczynowe, które także jest asymetryczne – tylko przyczyna wyjaśnia skutek, nie na odwrót.¹⁹

Odpowiedź, jaką na powyższą trudność zaproponował Hempel (1965, 352), można streścić następująco. Odróżniając prawa następstwa (uwzględniające zmianę w czasie) od praw współwystępowania (synchroniczna zależność funkcyjna między zmiennymi), stwierdza się wyjaśnienia przyczynowe tylko w przypadku tych pierwszych. Zależność między zmiennymi jest symetryczna tylko w przypadku praw współwystępowania, a te, zdaniem Hempela, nie są prawami przyczynowymi. Wyjaśnianie w koncepcji DN jest symetryczne tylko w przypadku praw współwystępowania, ale ponieważ te

załoby zastąpić przez bezkierunkowe zależności funkcyjne (klasyczny wyraz daje temu Russell 1912/1913), nie mogą odnieść się do tych faktów. Podobnie Hempela nomologiczno-dedukcyjny model wyjaśniania, gdzie konkretny przypadek wyjaśnia się przez subsumpcję pod prawo, nie potrafi zróżnicować takich przypadków”. (1998, 166-167)

¹⁸ Nie podejmuję tu systematycznej dyskusji na temat innej tezy o symetrii, mianowicie między wyjaśnianiem a przewidywaniem. Dotyczyć jej będą jedynie pewne uwagi w Punkcie 4.4 o kontrfaktycznym wnioskowaniu w bayesowskich modelach przyczynowych. Szczegółową analizę tej symetrii podaje J. Woleński 1979/1996. Konkluduje ją następująco: „symetria wyjaśniania i przewidywania jest istotną cechą metodologiczną systemów teoretycznych, a być może także kryterium odróżniającym systemy teoretyczne od wiedzy nieteoretycznej” (1979/1996, 265). Taka symetria – zdaniem S. Nowaka – przesądza o praktycznej doniosłości wyjaśnień przyczynowych (Nowak 1985, 354).

¹⁹ Por. (Metallmann 1934, 71; Kmita 1977, 188).

prawa nie są przyczynowe, więc i oparte na nich wyjaśnienia nie są przyczynowe. Symetrii wyjaśniania w koncepcji DN nie można więc uznać za przejaw nieadekwatności względem zależności przyczynowej, gdyż taka w przypadku praw współwystępowania nie zachodzi.

Jedno z praw dla gazów idealnych określa zależność ciśnienia i objętości niezależnie od czasu, jest więc prawem współwystępowania. Ale wiadomo, że to kompresja gazu spowodowała wzrost jego ciśnienia. G. H. von Wright (1973) w obronie koncepcji DN próbował uchylić tę trudność. Symetrię wyjaśniania można tu złamać dla konkretnego *zastosowania* funkcyjnej zależności określonej w prawie współwystępowania, gdy jedna ze zmiennych (np. objętość gazu czy wysokość pnia) jest poddana *manipulacji*. Skoro w danym przypadku dokonano manipulacji na jednej ze zmiennych, to tylko ona może wystąpić w opisie warunków wstępnych w eksplanansie, a pozostałe, zależne od niej funkcyjnie, będą występować w eksplanandum. Uzupełnienie koncepcji DN o manipulowalność pozwoli uniknąć trudności z nieadekwatnością względem asymetrii przyczynowej. Prowadzi to do następującej modyfikacji formuły (D):

wyjaśnić to wydedukować z prawa naukowego (D*M)

zastosowanego do zmiennej manipulowanej.

Pierwotne zamierzenie Hempla, jak odnotowałem wcześniej, dotyczyło wprowadzenia koncepcji wyjaśniania przyczynowego jako szczególnego przypadku wyjaśniania subsumpcyjnego wyrażonego w formule (D). Proponowana wersja tej formuły w postaci (D*M) zawiera pojęcie „manipulowalności”, które – wyeksplikowane lub nie – jest kategorią przyczynową. Tym samym (D*M) przekreśla perspektywę realizacji zadania sprowadzenia wyjaśniania przyczynowego do szczególnego rodzaju wyjaśniania subsumpcyjnego, gdzie byłoby ono czymś wtórnym, wyprowadzalnym w zasadzie z – bardziej „obiektywnej” – formy logicznej zdań tworzących wyjaśnienie.

Druga grupa zarzutów wykazuje, że koncepcja DN nie ustala warunku koniecznego wyjaśniania. W ostatecznej konkluzji, jak zwrócę na to uwagę, także one prowadzą do stwierdzenia, że podjęta przez Hempla redukcja koncepcji wyjaśniania przyczynowego nie była udana.

Tę grupę zarzutów zainicjował przykład M. Scrivena, ilustrujący fakt, że – przynajmniej niektóre – zadowalające wyjaśnienia nie muszą odwoływać się do praw naukowych, gdyż wystarczy, że określą przyczynę danego zda-

rzenia. I tak, moje przypadkowe uderzenie kolanem w biurko jest wyjaśnieniem tego, skąd wzięła się abstrakcyjnie rozbiegana plama atramentu na dywanie, mimo że nie podaję odnośnych praw Newtona czy innych (Scriven 1962, 90).

Oczywiście nie mamy tu do czynienia z sytuacją, gdzie nie potrafiąc podać pełnego i zadowalającego wyjaśnienia, proponujemy tylko określane przez Hempela jako „szkic wyjaśnienia”. Podane wyżej wyjaśnienie uznajemy za pełne (Scriven 1962, 94), twierdząc tym samym, że nie wszystkie wyjaśnienia muszą być nomologiczne, oparte na prawach naukowych. Stanowisko Scrivena, że istnieją wyjaśnienia konkretnych zdarzeń nieodwołujące się do praw naukowych, ma istotną presupozycję. Zależec będzie bowiem od metafizycznej tezy o istnieniu przyczynowości jednostkowej, a mianowicie o istnieniu zależności przyczynowych między konkretnymi zdarzeniami, których nie da się podporządkować zależnościom między odpowiadającymi im typami zdarzeń. Podany wyżej przykład musielibyśmy uzupełnić o (uzasadnioną na gruncie metafizyki przyczynowości) tezę, że moje uderzenie kolanem i strącenie kałamarza są w relacji, która nie jest szczególnym przypadkiem np. relacji, jakie zachodzą między uderzeniami powodującymi wstrząs biurka a spadaniem z niego przedmiotów.²⁰

Koncepcję, która uwzględnia wnioski z kontrprzykładu Scrivena, ale nie jego założenia metafizyczne, zaproponował w ostatnim czasie J. Woodward; szczegółowo omawiam ją w Punkcie 9.1.

Jako uzupełnienie koncepcji DN Hempel zaproponował koncepcję dedukcyjno-statystyczną (DS). Jeśli wśród przesłanek znajduje się nomologiczne twierdzenie, dotyczące regularności statystycznych (np. każdy atom C^{14} ma prawdopodobieństwo rozpadu w okresie 5 730 lat wynoszące 0,5), to otrzymujemy szczególny przypadek wnioskowania o schemacie DN, w którym wniosek-eksplanandum jest pewną regularnością statystyczną. Jak zaobserwował Salmon (1989, 53) i co jest potwierdzane przez innych (Psillos 2002, 242), koncepcja DS, która skupiała szczególną uwagę Hempela w (1965) jest szczególnym przypadkiem koncepcji DN zastosowanej do wyjaśniania regularności statystycznych. Z tego też powodu dziedziczy omówione wyżej trudności.

Zwróćmy w tym miejscu uwagę, że trudności koncepcji DN oraz DS związane są przede wszystkim z podjętą przez Hempela próbą

²⁰ Amsterdamski traktował przykład Scrivena jako kontrprzykład koncepcji wyjaśniania DN – a szerzej subsumpcyjnej – gdyż występuje tu wyjaśnienie przez przyczynę, ale bez odwołania się do prawa (1983, 121).

podporządkowania koncepcji wyjaśniania przyczynowego wyjaśnianiu subsumpcyjnemu. Krytyka tych koncepcji, którą tutaj przywołałem, nie podważa jednak pewnych elementarnych intuicji, jakie niezbyt adekwatnie wyrażają koncepcje DN i DS. W proponowanej niżej proceduralnej koncepcji wyjaśniania przyjmuję bowiem tę podstawową ideę, że wyjaśnienie polega na pokazaniu związku z pewną ogólną zależnością – w tym przypadku jest to model przyczynowy. Warunkiem jednak podania wyjaśnienia *przyczynowego* nie jest tylko związek logiczny między modelem czy opartymi na nim uogólnieniami projekcyjnymi a opisem wyjaśnianego zdarzenia. Warunkiem, który pominął Hempel, a który jest jednym z podstawowych w proponowanej niżej koncepcji proceduralnej wyjaśniania przyczynowego, jest spełnienie kryterium proceduralnego. To kryterium wymaga podjęcia określonych interakcji z badaną rzeczywistością, czego efektem jest otrzymanie danych manipulacyjnych o określonym wzorze zależności między zmiennymi. Ta podstawowa różnica w stosunku do koncepcji DN i DS pozwala, jak pokażę niżej, uniknąć przywołanych tu trudności koncepcji wyjaśniania, jakie zaproponował Hempel.

Hempel w (1965), uwzględniając niektóre z zarzutów wobec koncepcji DN oraz DS, wprowadził nową koncepcję wyjaśniania indukcyjno-statystyczną (IS), która ma dostarczać probabilistycznych wyjaśnień zdarzeń jednostkowych. W koncepcji IS założenie determinizmu przyczynowego zostaje uchylone przez Hempla.

Wyjaśnienie jednostkowego zdarzenia, np. tego, że Jan wyzdrowiał z zapalenia gardła wywołanego przez paciorkowca z grupy *Streptococcus* dlatego, że otrzymał penicylinę w zastrzykach, jest w koncepcji IS probabilistyczne. Jak podkreśla Hempel (1965, 381), nadal w wyjaśnieniu zwraca się uwagę na ogólną prawidłowość. W tym przypadku jednak nie można podać (deterministycznego) prawa przyrody, zgodnie z którym każdy z takim zapaleniem gardła, kto otrzyma penicylinę, wyzdrowieje. Ta zależność jest niedeterministyczna: zachodzi w „dużym procencie przypadków czy z wysokim prawdopodobieństwem statystycznym” (1965, 381). Po podaniu opisu warunków początkowych z tej zależności probabilistycznej można z *dużym prawdopodobieństwem* wywnioskować eksplanandum.

Być może wprowadzając wymóg dużego prawdopodobieństwa, Hempel spodziewał się, że w logice indukcji Carnapa zostanie szczegółowo opracowana reguła wnioskowania przypominającej regułę *modus ponens*: na podstawie tego, że jakieś zdanie jest bardzo prawdopodobne, można je uznać. Wobec wielu problemów sformułowanie takiej reguły okazało się jednak

zbyt trudne (Kawalec 2003a, 94-101), a wymóg wysokiego prawdopodobieństwa okazał się poważnym ograniczeniem koncepcji IS.

Bez określenia zależności przyczynowych, jakie są niezbędne do podania wyjaśnienia jednostkowego zdarzenia, samo podanie prawdopodobieństwa (regularności statystycznej), może być bardzo mylące. Najbardziej wyraźnym tego przykładem są przypadki występowania konfunderów: mimo dużego warunkowego prawdopodobieństwa zajścia jednego zdarzenia względem drugiego, oba mogą być warunkowo niezależne względem konfundera. Na tej obserwacji opiera się m.in. zarzut, jaki wobec koncepcji IS sformułował Salmon (1989, 58). Założmy, że jako wyjaśnienie faktu, że Jan wyzdrowiał ze zwykłego przeziębienia w ciągu tygodnia, podamy, iż przez ten czas przyjmował dużą dawkę witaminy C, np. 500 lub 1000 jednostek. Spełnione jest kryterium koncepcji IS, gdyż prawdopodobieństwo wyzdrowienia ze zwykłego przeziębienia wśród tych, którzy przyjmują zwiększoną dawkę witaminy C, jest wysokie. Jednak jeśli uwzględnić, że większość osób zdrowieje ze zwykłego przeziębienia w ciągu tygodnia (także ci, którzy nie przyjmują witaminy C), to podane wyjaśnienie wydaje się mylące, gdyż sugeruje zachodzenie zależności przyczynowej między wyzdrowieniem a przyjmowaniem witaminy C.

Koncepcja IS jest więc kolejną – dodajmy nieudaną – próbą redukcji koncepcji wyjaśniania przyczynowego do wyjaśniania subsumpcyjnego. W tym przypadku redukcja zasadzałaby się na sprowadzeniu zależności przyczynowych do wysokiego prawdopodobieństwa warunkowego. Taka jednak redukcja jest zbyt dużym uproszczeniem. Gdyby nawet zastąpić ją bardziej wyszukaną wersją, wprowadzając np. dyskutowane w Punkcie 4.4 założenie Markowa i założenie wierności, to nadal taka koncepcja wyjaśniania byłaby niezadowolająca. Od wyjaśniania przyczynowego żądam bowiem, aby spełniało kryterium proceduralne, którego nie da się zredukować do – uproszczonej czy wyszukanej – postaci wzorca obserwowanego wśród regularności statystycznych.

Poważny problem koncepcji IS, który także – jak pokażę w dalszych rozważaniach – występuje w późniejszych propozycjach koncepcji wyjaśniania przyczynowego, to określony przez Hempla problem dwuznaczności (1965, 394). W odróżnieniu od koncepcji DN czy DS, wprowadzenie dodatkowych informacji w eksplanansie wyjaśniania IS może prowadzić do tego, że w miejsce eksplanandum wyjaśniona zostanie (z dużym prawdopodobieństwem) jego negacja. Jeśli w poprzednim przykładzie taką dodatkową informacją będzie np. to, że Jan zaraził się odmianą paciorkowca

Streptococcus, która jest odporna na penicylinę, to jako eksplanandum, które jest bardzo prawdopodobne na podstawie eksplanansa, otrzymamy negację eksplanandum z wcześniejszego przykładu: jest bardzo prawdopodobne, że Jan nie wyzdrowieje (po podaniu penicyliny).

Problem dwuznaczności dotyczy więc tego, w jaki sposób określić klasę odniesienia, do której należy Jan. W zależności od klasy, do której zaliczymy Jana, prawdopodobieństwo zdarzenia opisywanego w eksplanandum może ulegać znacznym zmianom. W tym kontekście powraca więc założenie totalnego deskrypcjonizmu: określenie klasy odniesienia dla Jana powinno być maksymalnie specyficzne, a więc określone jak najbardziej szczegółowo tak, aby dodatkowe dookreślenia nie zmieniły prawdopodobieństwa eksplanandum.

Hempel uznał, że rozwiązaniem problemu dwuznaczności będzie takie określenie maksymalnie specyficznej klasy odniesienia, które wprowadza się z uwagi na całą dostępną w danym momencie wiedzę ogólną oraz naukową. Taka epistemiczna relatywizacja wydaje się niejednoznaczna (Psillos 2002, 251). W ten sposób bowiem nie jest wykluczony fakt znacznej rozbieżności między „obiektywnie” zdeterminowaną maksymalnie specyficzną klasą odniesienia, a wyodrębnioną przy relatywizacji epistemicznej. Ponadto otwiera to drogę do wprowadzenia założenia determinizmu przyczynowego, z uwagi na które koncepcja IS tylko wyraża niepełność wiedzy wymaganej do podania pełnego wyjaśnienia deterministycznego.

8.2 Zależności przyczynowe w wyjaśnianiu SR oraz CM

Kolejne koncepcje wyjaśniania projektowano, biorąc pod uwagę trudności koncepcji DN,²¹ zwłaszcza z adekwatnością w stosunku do zależności przyczynowych. Koncepcja SR (statystycznej relewancji) Salmona (1971b) jest najważniejszą z propozycji zgłoszonych po koncepcji Hempela. Podanie wyjaśnienia polega tu zasadniczo na określeniu, jakiego rodzaju cechy są statystycznie istotne dla eksplanandum, a jakie nie, oraz określeniu prawdopodobieństwa zdarzenia, uwzględniając wyodrębnioną jego charakterystykę

²¹ P. Kitcher i W. C. Salmon tak konkludują swoją analizę pragmatycznego modelu wyjaśniania B. van Fraassena (1977; 1980): „van Fraassen zaproponował piękną koncepcję pragmatyki wyjaśniania, która powinna być traktowana jako uzupełnienie raczej niż alternatywa w stosunku do tradycyjnych stanowisk w kwestii wyjaśniania” (1998, 188). Z tego też względu pragmatyczny model wyjaśniania van Fraassena nie będzie tu osobno dyskutowany.

jako istotną statystycznie. Zadowalające wyjaśnienie podaje prawdopodobieństwo eksplanandum E , które jest wyższe po uwzględnieniu wyodrębnionych jako statystycznie istotnych cech S , niż bez ich uwzględniania:

$$P(E) < P(E | S). \quad (8.1)$$

Podajmy uproszczony przykład wyjaśnienia w koncepcji SR. Pytamy, dlaczego dana osoba będąca nastolatkiem (N) jest przestępcą (R). Założmy, że jedyne cechy statystycznie relewantne dla K wśród N to płeć (K lub M) oraz miejsce zamieszkania (U – miasto, W – wieś). Otrzymujemy zatem partycję N na cztery możliwe klasy: NKU, NKW, NMU, NMW. Na wyjaśnienie składać się będą: (1) określenie prawdopodobieństwa bycia przestępcą dla całej klasy nastolatków; (2) oraz dla każdej z czterech wyróżnionych; (3) określenie, do której z nich należy ta osoba. W ten sposób uzyskamy informację, jaki zestaw cech okazał się statystycznie istotny dla wyjaśnianego faktu.²² Zakładając, że dana osoba należy do NMU, zadowalające wyjaśnienie musi spełniać warunek:

$$P(R) < P(R | NMU). \quad (8.2)$$

Zwróćmy uwagę zwłaszcza na dwa podstawowe założenia koncepcji SR. Jednym z nich jest koncepcja wyjaśniania przyczynowego. Drugie natomiast stwierdza, że relewancja statystyczna adekwatnie reprezentuje zależności przyczynowe.

Jeśli to założenie rozumieć w ten sposób, że na podstawie relewancji statystycznej można w każdym przypadku jednoznacznie określić zależności przyczynowe, to jest to zdanie fałszywe, co obszernie ilustruje wiele współczesnych prac (Pearl 2000; Spirtes, Glymour i Scheines 2000). Posłużę się przykładem Salmona. Ciśnienie atmosferyczne A powoduje występowanie

²² Kanoniczną postać wyjaśniania SR można scharakteryzować następująco. Pytanie o wyjaśnienie ma postać: Dlaczego x , będący członkiem klasy wyznaczonej przez cechę A , ma także cechę B ? Odpowiedź rozpoczyna się od wprowadzenia homogenicznej partycji A , czyli wyróżnienia klasy podzbiorów C_i zbioru A , które są wzajemnie rozłączne, a łącznie wyczerpujące. Partycja jest tak skonstruowana, że nie można uzyskać innej statystycznie relewantnej partycji żadnego z podzbiorów AC_i względem B , a więc nie ma innych cech D_k w A takich, że $P(B | AC_i) \neq P(B | AC_i D_k)$. Na właściwą część wyjaśnienia składają się: (1) określenie prawdopodobieństwa B w A , czyli $P(B | A) = p$; (2) określenie prawdopodobieństwa B dla każdego podzbioru A w homogenicznej partycji, czyli $P(B | AC_i) = p_i$ oraz (3) określenie, do której partycji należy x (Salmon 1971b, 76-77).

burz B oraz położenie strzałki S barometru. Oczywiście położenie strzałki nie jest przyczyną burzy, a więc S jest statystycznie nieistotne dla B , tzn. $P(B|AS) = P(B|A)$. Natomiast A jest relewantne dla B , gdyż $P(B|AS) \neq P(B|S)$. Tym samym Salmon uznaje A za istotne dla wyjaśnienia B , zaś S za nieistotne. Z tymi zależnościami statystycznymi zgodnych jest więcej struktur przyczynowych niż ta opisana przez Salmona. I tak, odwrócenie zależności przyczynowej między A a S daje dokładnie takie same zależności statystyczne, jak w przykładzie Salmona, ale S byłoby wówczas istotne w wyjaśnieniu B , przynajmniej jako przyczyna pośrednia.

Ta prosta modyfikacja przykładu Salmona pokazuje, że znajomość relewancji statystycznej nie zawsze pozwala na jednoznaczne określenie kierunku zależności przyczynowej, a tym samym na określenie, co mogłoby pełnić rolę eksplanansu – czy ciśnienie atmosferyczne, czy położenie strzałki barometru.

Tzw. paradoks Simpsona, gdzie obserwuje się zmianę prawdopodobieństwa warunkowego po zsumowaniu elementów danej partycji, pokazuje także, że kategoria relewancji statystycznej nie jest warunkiem koniecznym dla zależności przyczynowych. Okazuje się więc, że koncepcja SR dziedzi czy podobne do DN trudności z nieadekwatnością wobec asymetrii przyczynowej.

Salmon, uznając trudności koncepcji SR, zaproponował inną (1984). Mechanicystyczna koncepcja wyjaśniania przyczynowego (CM) (skrót ang. causal mechanical) zmierza do podania takiej charakterystyki zależności przyczynowych, która następnie mogłaby służyć do wyjaśniania zależności statystycznych, np. tego rodzaju, jakich wymagała koncepcja SR.

Zasadnicze pojęcie, które zostało tu wprowadzone, to *proces przyczynowy* rozumiany jako proces fizyczny, który może przenosić swoją strukturę w sposób ciągły, np. piłka (ale nie jej cień). Idea Salmona polegała na wprowadzeniu *fizycznego* związku między przyczyną a skutkiem, gdzie i przyczyna, i skutek są procesami przyczynowymi. Między procesami zachodzi lokalna jednorazowa *interakcja przyczynowa*, gdy jeden proces przyczynowy po zetknięciu z drugim pozostawia na nim taki fizyczny *ślad*, który jest dalej przekazywany,²³ np. zderzenie cząstek, które zmienia ich kierunek oraz energię kinetyczną. Przesuwający się cień jadącego samochodu (pseudoprocess) przy zetknięciu z cieniem innego obiektu także zmienia kształt (ślad), ale ta zmiana nie jest przekazywana – po minięciu go, cień

²³ Precyzyjną definicję Salmon wprowadza w (1984, 148).

samochodu wróci do pierwotnego kształtu. Proces przyczynowy jest więc procesem spełniającym jednocześnie dwa warunki: przenosi swoją strukturę oraz może przenosić taką modyfikację swojej struktury, która wystąpiła w jednorazowej lokalnej interakcji (1984, 144).

Wyjaśnienie danego zdarzenia polega w koncepcji CM przede wszystkim na określeniu (przynajmniej części) procesów przyczynowych i interakcji prowadzących do niego oraz opisaniu procesów i interakcji, które je tworzą. Na wyjaśnienie ruchu czerwonej bili i białej składa się to, że są one procesami przyczynowymi, podobnie jak kij bilardowy. Wyjaśnić to również podać interakcje, jakie między nimi zaszły, a więc uderzenie kija w białą bilę, a tej w czerwoną. Natomiast cień rzucany przez jedną z bil na drugą nie stanowi składowej wyjaśnienia, gdyż nie jest procesem przyczynowym.²⁴

Zasadniczy problem z tą propozycją Salmona polega na tym, że kategorie procesu, interakcji przyczynowej i przekazywanego śladu fizycznego są zbyt szerokie, by wyodrębnić te własności – w tym przypadku – kul bilardowych, które rzeczywiście są przyczynowo i ekspanacyjnie istotne, zwłaszcza ich masy i prędkości przed zderzeniem. Gdyby np. zdarzyło się tak, że niewielka ilość talku z kija pozostała na białej bili, a następnie przy uderzeniu część zostałaby na powierzchni bili czerwonej, to – zgodnie z podanymi przez Salmona kryteriami procesu przyczynowego – ten talk należałoby uznać za jeden z elementów wyjaśnienia CM.

Na inną trudność w koncepcji CM zwracają uwagę P. Dowe (2000, 71-74) oraz Psillos (2002, 113-18). Salmon zmierzał w kierunku redukcyjnej charakterystyki zależności przyczynowej, w której podstawowe dla jego koncepcji kategorie, jak proces przyczynowy, interakcja przyczynowa oraz ślad, byłyby zdefiniowane w innych kategoriach niż przyczynowe, np. stałego współwystępowania czy przecinania się procesów.

Nawet jednak podana w późniejszej pracy (Salmon 1997, 250) zmodyfikowana charakterystyka tych kategorii nie pozwala ani na taką redukcyjną analizę, ani nawet na uniknięcie kołowości związanej ze wzajemną definiowalnością interakcji przyczynowej i śladu. Przytoczę tutaj dwa istotne elementy tej charakterystyki:

S-III Ślad jest przekazywany w danym przedziale, gdy występuje w każdym punkcie czasoprzestrzeni tego przedziału, przy braku interakcji.

²⁴ Koncepcję pokrewną Salmona CM koncepcji wyjaśniania można znaleźć w (Krajewski 1976, 25-28).

S-IV Interakcja przyczynowa jest przecięciem, w którym w obu procesach występuje ślad (są zmienione) oraz ślad w każdym procesie jest przekazywany poza samo przecięcie.

W S-III nie można pominąć braku interakcji, gdyż w takim przypadku pseudoprocęs, w którym dany ślad byłby przekazywany przez (przypadkowe) zestawienie wielu interakcji, musiałby być uznany za proces przyczynowy, np. zmiana cienia, która następnie byłaby przekazywana przy jego przesuwanie się tylko dzięki przypadkowej konfiguracji przedmiotów, które po drodze mija przejeżdżający samochód.

Z kolei przekazywanie śladu, o którym mowa w S-IV, jest uznane przez Salmona za istotną charakterystykę procesów przyczynowych, która różni je od pseudoprocęśw. Nie można także zastąpić w S-III interakcji przez np. przecięcia, gdyż procesy przyczynowe przecinają się z wieloma pseudoprocęśwami. Po takim zastąpieniu znowu wiele procesów przyczynowych nie zostałyby zakwalifikowanych jako takie wedle zmodyfikowanej postaci warunku S-III.

Jak przyznał więc Salmon (1997a, 253), ślady nie są elementem konstytucji procesów przyczynowych, lecz jedynie ich symptomem.

Nawet gdyby udało się opracować strategię unikającą kołowatości w eksplikacji kategorii procesu przyczynowego, interakcji oraz śladu,²⁵ to nadal – jak zwraca uwagę Woodward – koncepcja wyjaśniania CM stałaby wobec poważnego problemu. Ona „przypisuje centralną rolę w wyjaśnianiu śledzeniu czasoprzestrzennie ciągłych procesów przyczynowych” (2003, 353), co w wielu przypadkach wyjaśnień nie ma zastosowania. Przykłady takich wyjaśnień, gdzie nie ma „śledzenia” procesów, to oddziaływanie na odległość lub przyczynowość przez brak lub podwójne zapobieganie, gdzie nie mamy do czynienia z występowaniem ciągłych czasoprzestrzennie procesów czy przekazywaniem energii od przyczyny do skutku. Ponadto w większości nauk społecznych wyjaśnienia nie odwołują się bezpośrednio do procesów fizycznych, które – jak być może twierdziłby Salmon – zachodzą na innej niż podana w wyjaśnieniu płaszczyzna. Żądanie, by w każdym przypadku śledzić procesy fizyczne, prowadziłoby, jak ilustruje Woodward (2003, 355), zarówno do pominięcia istotnych, jak i wprowadzenia zbędnych, informacji. Cena, jaką osiągają pomarańcze na rynku owoców w roku,

²⁵ Przyjmując za Dowem kategorię zachowania wielkości, Salmon podaje zarys takiego rozwiązania w (1997b). Krytykę tego rozwiązania formułują (Dowe 2000; Psillos 2002). Woodward (2003, 356-58) pokazuje, że powracają podstawowe trudności, które omówiłem w tekście w stosunku do kategorii śladu przenieszonego przez proces fizyczny.

w którym zmniejszyła się liczba zbiorów na skutek przymrozku, jest wyjaśniana przez krzywe popytu i podaży, ale nie przez szczegółowy opis fizycznego procesu np. opłaty dokonywanej za pomocą karty kredytowej przy zakupie pomarańcz.

8.3 Zależności przyczynowe w wyjaśnianiu jako unifikacji

Niekiedy próbuje się przywrócić koncepcję DN Hempela poprzez uzupełnienie jej jakimś dodatkowym elementem, uznając, że podaje ona adekwatne warunki konieczne. Takie uzupełnienie ma na celu uchylenie zarzutów, że ta koncepcja nie podaje warunku wystarczającego wyjaśniania.

Najważniejszą z takich propozycji jest koncepcja unifikacjonistyczna P. Kitchera.²⁶ Podstawową ideę wyraża on krótko:

Nauka prowadzi do postępu w naszym rozumieniu przyrody, pokazując nam, jak wyprowadzić opis wielu zjawisk, używając wciąż tego samego schematu wyprowadzania, a pokazując to, uczy nas, jak zredukować liczbę rodzajów faktów, które uznajemy za ostateczne. (1989, 432)

Przyjęta przez Kitchera modyfikacja wcześniejszych wersji koncepcji wyjaśniania (Friedman 1974) uzależnia unifikację nie od praw przyrody czy zbioru aksjomatów dedukcyjnej systematyzacji wiedzy u kresu nauki – co byłoby równoznaczne z uznaniem koncepcji regularnościowej praw – lecz od „schematów wyprowadzania”: „teoria unifikuje nasze przekonania, gdy dostarcza jednego (a ogólniej kilku) wzorca (wzorców) wnioskowania, który można użyć w wyprowadzaniu dużej liczby zdań przez nas uznawanych” (Kitcher 1981, 333).

Wyprowadzenie dedukcyjne z prawa naukowego – a szerzej, z uogólnienia o szerokim zakresie – Kitcher uznaje bowiem za warunek konieczny, lecz niewystarczający wyjaśnienia. To wyjaśniające uogólnienie musi ponadto unifikować bogaty zestaw zjawisk. Warunek ten spełniają wzorcowe przykłady, jak dokonana przez Newtona unifikacja teorii ruchu ciał niebieskich i ziemskich czy J. Maxwella unifikacja zjawisk elektryczności i magnetyzmu.

Kitcher nie podał ścisłych kryteriów dla zadowalającego wyjaśnienia, a jedynie w przybliżeniu określił, na czym ma polegać unifikacja. Im mniej różnych schematów wyjaśnienia używa się przy wyprowadzaniu możliwie

²⁶ Podstawowy zarys tej koncepcji podał M. Friedman (1974), a systematycznie rozwinął ją P. Kitcher (1989).

największej liczby różnorodnych zjawisk oraz im ściślejsze ograniczenia nakładają one na to, co można z nich wyprowadzić, tym bardziej zunifikowane jest zaproponowane wyjaśnienie. Pomimo że nie doprecyzowuje liczby wniosków i wzorców wyjaśnień ani ograniczeń, jak postąpić, gdy te żądania prowadziłyby w odmiennych kierunkach, to stwierdza, że w konkretnych przypadkach jest jasne, który z kandydatów na wyjaśnienie najlepiej spełnia te wymogi.

Czy długość cienia jest – w sensie unifikacjonistycznym – wyjaśnieniem wysokości pnia? Powszechnie przyjęte wzorce wyjaśniania w przypadku rozmiaru przedmiotów – takich jak np. góry, planety, organizmy czy artefakty – odwołują się do „warunków, w których taki obiekt powstał, oraz modyfikacji, którym następnie był poddany” (Kitcher 1989, 485). Dodanie do tych wzorców kolejnego, który pozwalałby wyjaśniać rozmiary przedmiotów w oparciu o cień, jaki rzucają, jest niezgodne, jak twierdzi Kitcher, z wymogami koncepcji unifikacjonistycznej.

Zgodnie z deklaracją Kitchera, koncepcja ta jest kontynuacją idei Hempla traktowania koncepcji wyjaśniania przyczynowego jako szczególnego przypadku wyjaśniania subsumpcyjnego. Kitcher (1989, 477) twierdzi, że zależność przyczynowa jest wtórna wobec zależności eksplanacyjnej. Niezależny od wyjaśniania porządek przyczynowy nie istnieje. Unifikacja eksplanacyjna nie zakłada żadnych pojęć przyczynowych, a określenie zależności przyczynowych opiera się na tej właśnie unifikacji. Jeżeli więc można mówić o asymetrii wyjaśniania związanej z zależnościami przyczynowymi, to dlatego, że takie asymetrie są wyznaczone przez (unifikacjonistyczną) koncepcję wyjaśniania.

Na problemy z wytłumaczeniem asymetrii przyczynowych na gruncie koncepcji unifikacjonistycznej zwrócił uwagę E. Barnes (1992). Podał on szereg przykładów asymetrii przyczynowych, gdzie trudno jest na gruncie koncepcji unifikacjonistycznej uniknąć nieintuicyjnego wniosku, że przyczyna jest wyjaśniana przez skutek.

Wysokość danego przedmiotu można wyprowadzić dedukcyjnie, znając położenie względem niego źródła światła oraz długość cienia. Takie wyprowadzenie nie jest wyjaśnieniem i – jak podkreśla Barnes – koncepcja unifikacjonistyczna powinna potwierdzić ten wniosek. W koncepcji Kitchera uwzględnia się klasę zdań, które mają być wyjaśnione, a które podpadają pod ten sam schemat: Przedmiot P ma długość d . Wyjaśnienie spełniające wymogi koncepcji unifikacjonistycznej odwoływać się będzie do powstania przedmiotu P i późniejszych jego modyfikacji. W przesłankach może więc

wystąpić opis intencji konstruktora danego przedmiotu (np. budynku), opis jego powstawania oraz trwałości tego przedmiotu i zachowania przez niego kształtu.

Przygodny fakt, że nie wszystkie przedmioty rzucają cień, nie może być – jak w koncepcji Kitchera – jedną z podstaw wyjaśniania. Wyjaśnianie przez cień nie podpada pod schemat wyjaśniania przez odwołanie się do powstania i modyfikacji przedmiotów, wobec czego jest innym wzorcem wyjaśniania. Można potraktować je albo jako wzorzec uzupełniający, albo jako zastępujący wyjaśnianie przez powstanie i modyfikację przedmiotów.

Jednak gdyby dodać wyjaśnianie przez cień jako nowy wzorzec, to w stosunku do już istniejących nie pozwalałby on na wyprowadzenie nowych wniosków. Ponadto stosowałyby się tylko do wąskiej klasy przedmiotów, które faktycznie rzucają cień. W stosunku więc do pozostałych (nieoświetlonych) przedmiotów wzorzec wyjaśniania przez cień musiałby zostać uzupełniony przez wzorzec wyjaśniania przez powstanie i modyfikację. Jednak ten ostatni wzorzec obejmuje przypadki wyjaśniane przez cień, w związku z czym otrzymalibyśmy unifikację wiedzy zawierającą oba wzorce, która byłaby gorsza od unifikacji zawierającej tylko wzorzec wyjaśniania przez powstanie i modyfikację.

Gdyby natomiast wzorzec wyprowadzania kształtu przedmiotu na podstawie jego cienia miał zastąpić dotychczas przyjęte wzorce wyjaśniania (przez powstanie i modyfikacje przedmiotów), to znacznie zmniejszyłaby się liczba przypadków, które można byłoby wyjaśnić do tych tylko, które rzucają cień. Wystarczy zwrócić uwagę choćby na przedmioty przezroczyste, które nie rzucają cienia i w stosunku do których wzorzec wyjaśniania przez cień nie stosuje się w postaci niezmodyfikowanej. W obu więc przypadkach okazuje się, że wzorzec odwołujący się do cienia jest mniej unifikujący niż dotychczas przyjęty. Cień pnia w sensie unifikacjonistycznym nie wyjaśnia jego wysokości.

Nie można więc wzorca odwołującego się do cienia przedmiotu ani dodać do dotychczas przyjętych, ani wprowadzić go na miejsce wzorca odwołującego się do powstania i modyfikacji. Tym samym asymetria wyjaśniania (długość wyjaśnienia cień, ale nie na odwrót) byłaby zachowana w koncepcji unifikacjonistycznej.

Co byłoby jednak – zapytajmy za Barnesem – gdyby wszystkie przedmioty rzucały wystarczająco dużo cieni, aby określić wszystkie ich wymiary? W takim przypadku moc unifikująca obu wzorców wyjaśniania:

przez cień oraz przez powstanie i modyfikację przedmiotów, jest taka sama. Asymetria wyjaśniania znowu staje się problemem.

W dyskusji koncepcji unifikacjonistycznej Kitcher jest świadomy takiego problemu i dyskutując go, wprowadza predykat wyrażający dyspozycję przedmiotu do rzucania cienia (1989, 486-87). Następnie odrzuca taką możliwość z uwagi na to, że ten predykat jest nieprojekcyjny, oraz wymóg, by ostateczna unifikacja wiedzy naukowej nie zawierała predykatów nieprojekcyjnych. W takim jednak przypadku, jak podkreśla Barnes, stajemy wobec pewnego konfliktu w poglądach Kitchera:

Przez „predykat projekcyjny” Kitcher wydaje się rozumieć coś w rodzaju predykatu odnoszącego się do rodzaju naturalnego, chociaż nie daje żadnej wskazówki, jak pogodzić realizm w kwestii rodzajów naturalnych [...] z przy czynowym antyrealizmem, czyli tezą należącą do grona najważniejszych składników jego teorii. (Barnes 1992, 564)

Przyjęcie poglądu, który Barnes określa realizmem w kwestii rodzajów naturalnych, jest niezgodny z podstawową ideą, którą Kitcher przejął od Hempla, a mianowicie traktowania wyjaśniania przyczynowego jako szczególnego przypadku wyjaśniania subsumpcyjnego. Jest także bardzo problematyczny dla ogólnego neo-Hume’owskiego programu, jaki stara się zrealizować Kitcher, proponując koncepcję unifikacjonistyczną.

Barnes podaje także przykłady wzorców wyjaśniania, które na gruncie koncepcji unifikacjonistycznej mają równą moc unifikacyjną, jednak nie zachowują asymetrii wyjaśniania. Te przykłady są tak dobrane, by zablokować inną strategię argumentacyjną, jaką w przykładzie z cieniem mógł stosować Kitcher. Ta strategia to poszerzanie dziedziny stosowalności danego wzorca. W przypadku cienia Kitcher mógł twierdzić, że wzorzec wyjaśniania przez powstanie i modyfikację przedmiotów stosuje się nie tylko do ich długości, jak wzorzec wyjaśniania przez cień, ale i do innych wymiarów, a nawet do struktury wewnętrznej.

Ta strategia poszerzania nie ma jednak zastosowania do przykładu symetrycznych względem czasu systemów zamkniętych, które podaje Barnes jako kolejny przykład (1992, 564-66). Może być to zamknięty system cząstek S , w którym obowiązują symetryczne względem czasu prawa Newtona. Mając pełny opis położenia i prędkości cząstek w tym systemie w danym czasie t , można – posługując się prawami Newtona – podać wyjaśnienie wszystkich zdań podpadających pod schemat: „Przedmiot P w S ma położenie L oraz prędkość V w czasie T ”. Wzorzec wyjaśniania prospektywny, gdzie $T = t^* > t$ ma tę samą moc unifikacyjną co wzorzec wyjaśniania retro-

dyktywny, gdzie $T = t^* < t$, mimo że nie jest on – intuicyjnie rzecz ujmując – wyjaśniający.

W tym przypadku Kitcher nie może dla zachowania asymetrii wyjaśniania zastosować strategii poszerzania. Do każdej dziedziny spełniającej wymogi zastosowań wzorca prospektywnego można bowiem zastosować – bez utraty mocy unfikacyjnej – koncepcję retrospektywną.

Przyznanie jednak, że mają taką samą moc unifikacyjną, oznaczałoby akceptację symetrii między stanami wcześniejszymi a późniejszymi opisywanego tu systemu *S*. Ostatecznie więc uniemożliwiłoby to wprowadzenie asymetrii między przyczyną a skutkiem w koncepcji unifikacjonistycznej, a tym samym podważałoby ideę, że wyjaśnianie przyczynowe jest szczególnym przypadkiem wyjaśniania subsumpcyjnego.

Powtórzmy więc za Barnesem:

Braki obecne w teorii [wyjaśniania] Kitchera odzwierciedlają braki modelu DN Hempla i po raz kolejny ujawniają istotny wymiar przyczynowy wyjaśniania naukowego. (1992, 570)

Woodward na kilka sposobów tłumaczy źródła problemów unifikacjonistycznej koncepcji wyjaśniania (2003, 362-67). Unifikacja, jak pokazują przykłady podawane przez Barnesę, nie ujmuje adekwatnie asymetrii wynikających z zależności przyczynowych, a więc, jak zwraca uwagę Woodward, asymetrii związanych z rzeczywistością lub kontrfaktyczną manipulacją przedmiotami. Jeśli zmieniłbym długość kołka, to także zmieniałby się długość jego cienia, ale nie na odwrót.

Tę nieadekwatność kategorii unifikacji można z kolei wytłumaczyć jej niejednorodnością. Niekiedy o unifikacji mówi się wtedy, gdy stworzony zostaje nowy schemat klasyfikacyjny lub zasób słownictwa, jak np. w opracowanej przez Linneusza klasyfikacji biologicznej. Inny rodzaj unifikacji to opracowanie formalizmu matematycznego, który można stosować do różnych rodzajów zjawisk, jak np. systemy równań opracowane przez Lagrange'a i W. Hamiltona. Inny jeszcze rodzaj unifikacji Woodward określa jako unifikację fizyczną, gdzie w przypadku zjawisk uważanych za niezwiązane ze sobą „pokazuje się, że są rezultatem wspólnego zbioru mechanizmów lub związków przyczynowych” (2003, 362), czego przykładem jest teoria grawitacji Newtona. Tylko w przypadku tej ostatniej unifikacji istnieje związek z zależnościami przyczynowymi. Osiągnięcie unifikacji w pierwszym i drugim sensie nie prowadzi do ustalenia żadnych bezpośrednich związków przyczynowych.

Ponadto przyjęcie określonych unifikujących schematów klasyfikacyjnych czy unifikującego formalizmu matematycznego może nie mieć nic wspólnego z ich mocą eksplanacyjną. Mogą być to bowiem ich własności pragmatyczne, np. prostota, łatwość użycia, powszechność stosowania. Przyjęte więc przez Kitchera pojęcie unifikacji nie różnicuje między przypadkami, kiedy ma lub nie ma ono związku z mocą eksplanacyjną. Ostatecznie unifikacja w ujęciu Kitchera jest pojęciem ekonomii opisu lub kompresji informacji – wiedza zunifikowana pozwala na wywnioskowanie jak najwięcej z jak najmniejszej liczby założeń lub przy jak najmniejszej liczbie schematów wyjaśniania. Zunifikowany opis, który nie zawiera informacji o zależnościach przyczynowych, występuje w licznych przykładach w nauce, np. w klasyfikacjach biologicznych, geologicznych i astronomicznych. Jeśli wiem, że dany organizm należy do określonej kategorii klasyfikacyjnej, np. jest ssakiem, to mogę na tej podstawie wiele wywnioskować na jego temat, np. że ma kręgosłup, serce, że jest żyworodny itd. Tego typu schematy zostaną jednak uznane przez naukowców za opisowe,²⁷ gdyż nie stanowią wyjaśnienia, np. nie podają przyczyn wyjaśniających, dlaczego ten organizm ma serce.

Innym, obok klasyfikacji, przykładem zunifikowanego ujęcia informacji opisowych w sposób ekonomiczny są różnorodne procedury statystyczne, np. analiza czynnikowa czy techniki skalowania wielowymiarowego. Na przykład, znając „obciążenie” każdego z n testów na inteligencję, można wyprowadzić $n \cdot (n-1) / 2$ wniosków o wzajemnych korelacjach tych testów (Woodward 2002, 49). Taka unifikacja nic jednak nie mówi o zależnościach przyczynowych.

Woodward, mimo licznych porównań koncepcji manipulacyjnej wyjaśniania z unifikacjonistyczną, nie ma gotowej odpowiedzi na pytanie, jak w swojej koncepcji zrekonstruować podstawowe intuicyjne idee związane z wyjaśnianiem unifikacjonistycznym. Bardziej szczegółowo tę uwagę rozwijam w kolejnym punkcie, poświęconym tej właśnie koncepcji wyjaśniania. Zaznaczę tu tylko krótko, że zasadniczym źródłem tego problemu jest pojęcie interwencji (por. Punkt 7.3). Warunki interwencji nie nakładają wymogu podjęcia faktycznej interakcji z badaną rzeczywistością, w związku z czym w koncepcji manipulacyjnej nie ma czynnika stymulującego osiągnięcie wyjaśnienia jak najszerszego zakresu zjawisk.

²⁷ W odniesieniu do nauk społecznych to stwierdzenie szczegółowo uzasadnia M. Styczeń (1977, 141-143).

Projekcyjność w wyjaśnianiu przyczynowym

9.1 Projekcyjność w manipulacyjnej koncepcji wyjaśniania

Współcześnie dość powszechne wśród filozofów nauki i metodologów jest przekonanie, że główny problem dotychczasowych koncepcji wyjaśniania tkwi w ich nieadekwatności względem zależności przyczynowych. Tę nieadekwatność można rozumieć różnorako. W proponowanym tu podejściu proceduralnym zasadnicza nieadekwatność dotychczas proponowanych koncepcji wyjaśniania polega na nieodróżnianiu zadania identyfikacji zależności przyczynowych od kwestii filozoficznej eksplikacji ich znaczenia. Konsekwencjami tego mieszania zadań, jakie wskazałem w najważniejszych współczesnych koncepcjach wyjaśniania, była bądź próba redukcji wyjaśniania przyczynowego do wyjaśniania subsumpcyjnego, bądź niezadawalające próby podania nieredukcyjnej koncepcji wyjaśniania przyczynowego.

Ze współczesnych najbliższą proponowanej w tej pracy koncepcji wyjaśniania jest koncepcja manipulacyjna Woodwarda. Wyjaśnianie w nauce jest uznane za asymetryczne, czego źródłem są zależności przyczynowe. Podanie zależności przyczynowych wiąże się z wyliczeniem związanych z wyjaśnianym faktem uogólnień, które są niezmiennicze w stosunku do pewnej klasy interwencji – faktycznych lub możliwych (kontrafaktycznych) – na wyjaśnianą rzeczywistość.¹ Poza tym zakresem przytaczane w wyjaśnieniu niezmiennicze uogólnienia mogą zupełnie tracić swoją obowiązywalność. Te uogólnienia też mogą prowadzić tylko do wyjaśnień probabilistycznych, a nie deterministycznych. Wyjaśnianie poprzez prawa przyrody stanowi

¹ Podobną własność uogólnień podkreśla w *Prawach nauki* W. Krajewski (1998, 48-49).

ważną grupę, ale nie wyczerpuje wszystkich rodzajów wyjaśniania, podobnie jak wyprowadzanie dedukcyjne nie jest warunkiem koniecznym wyjaśniania:

wyjaśnianie to sprawa podania systematycznych wzorców zależności kontrfaktycznych. Uogólnienia [...] nie tylko można wykorzystać do pokazania, że eksplananda [...] były czymś spodziewanym, wzięwszy pod uwagę faktycznie zaistniałe warunki początkowe i brzegowe, ale można je także wykorzystać do pokazania, jak te eksplananda *zmieniłyby się*, gdyby te warunki początkowe i brzegowe zmieniały się w różny sposób. [...] To właśnie ten rodzaj informacji umożliwi nam dostrzeżenie, że (i jak) warunki wymienione w eksplanansie [...] są eksplanacyjnie istotne dla tych eksplanandów. Są to także informacje istotne dla manipulacji i kontroli zjawisk opisywanych przez te eksplananda. (Woodward 2003, 191)

Podkreślmy, że dla koncepcji manipulacyjnej istotna jest podklasa zdań kontrfaktycznych:² „za istotny rodzaj zdań kontrfaktycznych uznają te, które opisują rezultaty interwencji” (2003, 196).

Zdaniem Woodwarda, źródłem wielu z omówionych wcześniej problemów dotychczasowych koncepcji wyjaśniania jest to, że nie „uwzględniają schematu kontrfaktycznej zależności między eksplanansiem a eksplanandum” (2002, 50). Ponadto

zwykliśmy sądzić, że wyjaśnianie ma coś wspólnego z określeniem zależności przyczynowych, a nie da się zaprzeczyć, że między nimi a zależnościami kontrfaktycznymi zachodzi ścisły związek. Jeśli wyjaśnienia odwołują się do przyczyn, to wydaje się bardzo prawdopodobne, że jakaś postać kontrfaktycznej teorii wyjaśniania musi być poprawna. (2000, 210)

Zmiana wysokości pnia (przy pozostawieniu pozostałych czynników przyczynowych bez zmian) wydłuży lub skróci długość jego cienia. Manipulacje długością cienia natomiast, które nie naruszają bezpośrednio samego pnia, nie doprowadzą do zmiany jego wysokości. Długość cienia jest więc kontrfaktycznie zależna od wysokości pnia, a tym samym jest też przez nią wyjaśniona. Oczywiście, odwrotna zależność nie zachodzi, gdyż z dwóch możliwych tu zdań wyrażających zależność kontrfaktyczną, prawdziwe jest tylko jedno: „Gdyby zmienić wysokość pnia, zmieniłaby się długość jego

² Kontrfaktyczne teorie przyczynowości rozpowszechniły się we współczesnej filozofii, a także metodologii nauk pod wpływem D. Lewisa (1973), który zainicjował także opracowanie formalnej semantyki zdań kontrfaktycznych.

cienia”. Stąd wynika także asymetria wyjaśniania, która odzwierciedla asymetrię kontrfaktycznej zależności przyczynowej.³

Jedną z istotnych konsekwencji tej koncepcji ma być legitymizacja wyjaśniania w naukach społecznych, gdzie – zdaniem Woodwarda – można co najwyżej podać uogólnienia,⁴ które są niezmiennicze względem pewnej klasy interwencji. Inną konsekwencją manipulacyjnej koncepcji wyjaśniania jest wprowadzenie stopniowości odróżnienia między przypadkowymi uogólnieniami a prawami przyrody w miejsce tradycyjnie przyjmowanej binarności. Zasadnicza różnica w tej koncepcji to zakres klasy niezmienniczości.

Zmierzając do wyartykułowania manipulacyjnej koncepcji wyjaśniania, Woodward krytykuje następujący przykład: Wszystkie kruki są czarne, k jest krukem, a zatem k jest czarny. Zdaniem Woodwarda, koncepcja DN potraktuje to wnioskowanie jako wyjaśnienie, gdyż eksplanans zawiera prawo, z którego logicznie wynika wniosek. Koncepcja ta nie spełnia ona jednak wymogów koncepcji manipulacyjnej, gdyż

zawiera warunek nomologicznie wystarczający dla rezultatu, ale pojęcie zmienienia tego warunku nie jest dobrze zdefiniowane lub wyjaśnienie nie opisuje poprawnie, jak zmieniłyby się ten rezultat przy zmianach wprowadzonych w tym warunku. (2003, 193-94)

A zatem Woodward odrzuca przyjęte przez Hempla traktowanie wyjaśniania przyczynowego jako szczególnego rodzaju wyjaśniania subsumpcyjnego. Dodatkowym powodem, który omawiam niżej, jest niekonieczność związku wynikania logicznego, który łączy eksplanans z eksplanandum.

W jakiej jednak mierze przykłady, jak podany wyżej dla kruków, podważają koncepcję Hempla? Przyjmując, że najbliższa Hempłowi byłaby regularnościowa koncepcja praw najlepszego systemu, można podać Woodwardowi następującą odpowiedź. Uogólnienie „Wszystkie kruki są czarne” nie

³ Występujący tu implicite warunek nakładany na zdania kontrfaktyczne zawarte w wyjaśnieniach Woodward wyraża następująco: „zdania kontrfaktyczne istotne dla wyjaśniania mają poprzedniki, które stają się prawdziwe dzięki specjalnemu rodzajowi egzogenicznego procesu przyczynowego, który nazywam interwencją. [...] interwencje można traktować jak manipulacje, które człowiek mógłby przeprowadzić w wyidealizowanych warunkach eksperymentalnych” (1997, S29). To, czy jest możliwa taka ludzka interwencja, jest faktem przygodnym i nie przesądza ona o związku przyczynowości z wyjaśnianiem. Por. też (Nowak 1985, 327-339; Brzeziński 1996, 280-286).

⁴ Zwracając uwagę na trudności w matematyzowaniu tez nauk społecznych, podobną tezę podał A. Szymański (1996, 131-33).

spełnia warunków, jakie nakładane są na prawa naukowe, gdyż nie można oczekiwać, że będzie ono elementem zbioru najlepszych systematyzacji dedukcyjnej wiedzy naukowej u kresu. W takiej systematyzacji kruki zostałyby scharakteryzowane w bardziej podstawowych kategoriach biologicznych, natomiast kolor jako względna własność fizyczna budzi poważne zastrzeżenia co do włączenia tego uogólnienia do zbioru podstawowych uogólnień naukowych. Tym samym więc nie ma podstaw, by przypisywać Hempelowi ten przykład jako trafną ilustrację wyjaśniania w koncepcji DN, gdzie – przypomnijmy – eksplanans musi zawierać *prawa*, a nie dowolne uogólnienia.

Krytykę Woodwarda należałoby zatem odnieść do koncepcji praw, jaka stoi u podstaw DN. Przyjmując, że jest to koncepcja regularnościowa, krytykę tę można byłoby wysłowić następująco. Koncepcja wyjaśniania DN prowadzi do uznania za wyjaśnienie każdego wniosku, w którym jako przesłanki występują prawa przyrody. Określenie, że są to składowe najlepszej dedukcyjnej systematyzacji wiedzy naukowej u kresu, nie gwarantuje, że określony w nich jest związek z kontrfaktycznym opisem, jak zmieniać się będzie wyjaśniane zjawisko wskutek podjętych interwencji. Zgodnie z przytoczonym wcześniej rozumowaniem Woodwarda, zachodzenie związku eksplanacyjnego wymaga określenia takiego związku. Dochodzimy więc do stwierdzenia możliwości, że prawa przyrody wyodrębnione przez najlepsze systematyzacje wiedzy naukowej, nie będą miały mocy eksplanacyjnej w sensie wskazanym przez Woodwarda. Żądanie w koncepcji DN przytaczania takich praw może więc prowadzić do sytuacji, gdzie mimo spełnienia wymogów tej koncepcji nie będziemy mieli do czynienia z wyjaśnieniem.

Zwróćmy w tym kontekście uwagę na dwa wymogi, jakie Hempel stawia eksplanansowi. Pierwszym jest kryterium syntaktyczne, by prawa występujące w eksplanansie „posiadały treść empiryczną”, czyli implikowały testowalne konsekwencje (1948, 137). Drugim warunkiem jest prawdziwość, którą Hempel uznaje za kryterium empiryczne. Możliwe jest takie odczytanie tych wymogów, przy którym w interesującej klasie przypadków pokrywają się one z wymogami proponowanymi przez Woodwarda.

Koncepcja manipulacyjna wyjaśniania nakłada wymóg określenia związku między eksplanansem a zmianami, którym może podlegać eksplanandum wskutek podejmowanych na nim interwencji. Zwróćmy jednak uwagę, że taki wymóg jako eksplanandum wprowadza nie jedno zdarzenie, lecz klasę zdarzeń powiązanych kontrfaktycznie z uwagi na pewne czynności dotyczące tych zdarzeń. Dla tak poszerzonego eksplanandum warunek

testowalności empirycznej Hempela może prowadzić do żądanych zmian. Można bowiem, zgodnie z tym warunkiem, zażądać, by wprowadzone w eksplanansie prawo posiadało konsekwencje dla poszerzonego eksplanandum względem określonej klasy czynności względem eksplanandum.

W stosunku do żądanej kontrfaktycznej relewancji wymóg prawdziwości, jaki na eksplanans nakłada Hempel, może wydawać się pewnym ograniczeniem. Zgodnie bowiem z analizą zaproponowaną w Punkcie 7.3 pojęcie interwencji Woodwarda może prowadzić do określania takich zależności kontrfaktycznych, którym nie odpowiadają żadne regularności. interesująca klasa przypadków dotyczy wyłącznie takich sytuacji, gdzie uwzględnia się tylko zaobserwowane regularności i w obrębie tej klasy, jak starałem się pokazać, nie ma zasadniczej rozbieżności między wymogiem relewancji eksplanacyjnej w wersji zaproponowanej przez Woodwarda a warunkami adekwatności dla koncepcji DN uzupełnionej o regularnościową koncepcję praw.

Oczywiście, dyskusja między obydwoma koncepcjami wyjaśniania mogłaby przenieść się na grunt koncepcji praw przyrody. Przypomnę, że propozycja Woodwarda polega na potraktowaniu praw jako szczególnej kategorii uogólnień niezmienniczych. Ich cechą istotną jest to, że nie tylko zdają sprawę z zaistniałych regularności (tego, co się zdarza lub zdarzyło), lecz pozwalają określić, co by się zdarzyło, gdyby podjąć określone interwencje. W przypadku wspomnianego wcześniej prawa dla gazu idealnego można powiedzieć, że pełni ono istotną rolę w wyjaśnianiu nie dlatego, że wzrost objętości gazu traktuje się jako nomicznie spodziewany, ale dlatego, że pozwala określić, jak zmieniłoby się jego ciśnienie, gdyby zmieniły się warunki. Jak podkreśla Woodward, ważnym elementem takiego wyjaśniania jest więc umieszczenie eksplanandum „w przestrzeni możliwych alternatyw” (2000, 209).⁵

Ta przestrzeń możliwych alternatyw może być jednak określona, jak wynika z przeprowadzonych w Punkcie 7.3 analiz manipulacyjnej koncepcji praw, w sposób niekołowy w stosunku do interwencji jako wyidealizowanego eksperymentu tylko poprzez odwołanie się do kategorii metafizycznych, jak mechanizm nomologiczny czy możliwości. Zwolennik regularnościowej koncepcji praw może jednak postawić pytanie o kryteria wyodrębnienia przestrzeni możliwych alternatyw relewantnych w stosunku do danego eksplanandum. Jeżeli byłaby nią określona regularność, wówczas

⁵ To stwierdzenie jest zbieżne z koncepcją wyjaśniania jako subsumpcji proceduralnej, którą prezentuję w Punkcie 9.3.

otrzymalibyśmy szczególny przypadek koncepcji DN, gdzie w miejsce eksplanandum proponuje się nie jednostkowe zdarzenie, lecz regularność.

Być może jednak przestrzeń możliwych alternatyw jest czymś więcej. Być może są to zdarzenia powiązane z eksplanandum modalnie. W takim jednak przypadku powracamy do prezentowanych w Rozdziale 7. trudności związanych z rozwiązaniem problemu identyfikacji i inferencji. W jaki sposób wyodrębnić tę klasę? Jak eksplanandum wynika z określonej modalnie zależności między zdarzeniami tworzącymi tę klasę?

Modyfikując na obecne potrzeby zarzut Psillosa (2002, 105-106), można powiedzieć, że Woodward, ignorując odróżnienie problemu eksplikacji zależności przyczynowej od jej identyfikacji, tworzy hybrydę pojęciową. Traktuje bowiem pewną postać manifestacji zależności przyczynowej jako jedną z jej podstaw pojęciowych. Tym samym otrzymujemy koncepcję, która nie pozwala na identyfikowanie zależności przyczynowych i daje kłopotliwą eksplikację tejsze zależności.

Co w przypadku, gdy eksplanans nie zawiera praw? Woodward jako wystarczający warunek wyjaśnienia uznaje przytoczenie takiego uogólnienia, które jest niezmiennicze względem pewnej klasy interwencji na zjawisko opisywane w eksplanandum. Może być to uogólnienie, które dopuszcza wyjątki poza tą klasą interwencji i w związku z tym nie kwalifikuje się na prawo przyrody (pomijając już choćby takie kryteria, jakie na prawa nakłada koncepcja najlepszego systemu). Byłby to przypadek wyjaśnienia, które byłoby jednak odrzucone na gruncie koncepcji DN. Można jednak podnieść kilka wątpliwości co do istnienia takiej różnicy.

Po pierwsze, Woodward nie podaje argumentu, że takie nienomologiczne wyjaśnienia nie są tylko tymczasowe. Otwarta jest przecież dla zwolennika DN możliwość argumentowania, że kiedy osiągnięta zostanie dedukcyjna systematyzacja nauki u kresu, przypadki wyjaśnień nomologicznych wyginą, a ich miejsce zajmą wyjaśnienia spełniające wymogi DN.

Po drugie, zwolennik koncepcji DN mógłby próbować pokazać, że nie potrafimy wyodrębnić przypadków wyjaśnień nienomologicznych. Nie ma jasnych kryteriów uogólnień niezmienniczych, które nie są prawami przyrody. Podana przez Woodwarda charakterystyka tej klasy uogólnień opiera się na tym, że dopuszczają one wyjątki poza obszarem niezmienniczości z uwagi na pewną klasę interwencji. Ten sam jednak warunek spełniają przypadkowe uogólnienia, które także są niezmiennicze z uwagi na pewną klasę interwencji. Dla Woodwarda jest to różnica tylko stopnia. Dla zwolennika koncepcji DN jest to niejasna płynność między tymi klasami uogólnień,

która pokazuje, że poza wyjaśnieniami nomologicznymi nieznane są inne wyjaśnienia.

Dochodzimy zatem do wniosku, że zwolennik DN może pokazać, iż manipulacyjna koncepcja przyczynowości jest jedynie artykulacją pewnych elementów koncepcji DN, a poza interesującą klasą przypadków te dwie koncepcje wyjaśniania w zasadzie się nie różnią.

Wskazany w powyższej argumentacji przeze mnie powodem tych trudności koncepcji manipulacyjnej jest kontrfaktyczna charakterystyka interwencji oraz przyjęcie uogólnień niezmienniczych jako podstawy wyjaśniania. Powrócę do tych dwóch źródeł trudności koncepcji Woodwarda poniżej, proponując w obu przypadkach inne rozwiązania: potraktowanie kontrfaktyczności jako pojęcia wtórnego wobec przyczynowości oraz wprowadzenie uogólnień projekcyjnych w miejsce niezmienniczych.

Koncepcja zaproponowana przez Woodwarda⁶ prowadzi do przemyślenia na nowo roli wynikania logicznego w wyjaśnianiu. Hempel w swojej koncepcji eksponował wyprowadzenie dedukcyjne eksplanandum z prawa naukowego, gdyż to wyprowadzenie miało pokazać, dlaczego należało się spodziewać zdarzenia, o którym mowa w eksplanandum. W Woodwarda koncepcji manipulacyjnej wyjaśniania wyprowadzenie dedukcyjne z prawa naukowego jest, jak podkreśla (2003, 201), niekiedy tylko skutecznym sposobem przekazania informacji o zależności kontrfaktycznej, a tym samym podania wyjaśnienia. Wyprowadzenie wysokości pnia z długości cienia w oparciu o prawa optyki geometrycznej jest przykładem na to, że nie wszystkie wyprowadzenia zgodne ze schematem wprowadzonym przez Hempla w koncepcji DN podają zależność kontrfaktyczną. Zatem schemat wyprowadzenia eksplanandum określony w tej koncepcji nie jest warunkiem koniecznym zależności kontrfaktycznej.

Od wyjaśnień Woodward żąda, by określały zależności pozwalające odpowiedzieć na pytanie, jak zmieniłoby się eksplanandum, gdyby zaszły takie to a takie zmiany w warunkach początkowych czy brzegowych. Wyprowadzenie dedukcyjne z zależności nomologicznej spełnia ten warunek, gdy „śledzi lub odzwierciedla relacje zależności fizycznej, które zachodzą między warunkami podanymi w eksplanansie a zjawiskami opisywanymi w eksplanandum” (2003, 201). Są jednak inne niż wynikanie logiczne z zależności nomologicznej sposoby przekazania takiej informacji, która ma moc eksplanacyjną. Chodzi zwłaszcza o

⁶ Zasadnicze elementy tej koncepcji wyjaśniania można odczytać też w pracach S. Nowaka, zwł. (1985, 351-381).

Inne sposoby reprezentacji, takie jak diagramy oraz grafy, [które] w podobny sposób przekazują informacje o zależności przyczynowej bez wyraźnego wprowadzania wyprowadzeń dedukcyjnych. (2002, 51)

Inne jeszcze sposoby ukazania zależności kontrfaktycznej niż dedukcyjne wyprowadzenie z prawa naukowego występują w przypadku przyczyn zdarzeń jednostkowych. I tak moje uderzenie kolanem w biurko jest wyjaśnieniem strącenia kałamarni, gdyż zachodzi w tym przypadku (asymetryczna) zależność kontrfaktyczna.

Do roli wynikania logicznego między eksplanansem a eksplanandum wrócimy jeszcze, dyskutując proponowaną tu proceduralną koncepcję wyjaśniania. Walentność zachowują uwagi Woodwarda dotyczące tego, że nie jest to ani warunek konieczny, ani wystarczający wyjaśniania. Jego koncepcja manipulacyjna nie zawiera dostatecznej charakterystyki konceptualnej zależności między eksplanandum a eksplanansem. Ograniczenie się do ogólnego stwierdzenia, że taka zależność powinna „śledzić lub odzwierciedlać relacje zależności”, jest dalece niewystarczające. Podobnie niewystarczające są pobieżne uwagi Woodwarda, że

nie można wyczytać z samych logicznych czy dedukcyjnych relacji pomiędzy rzekomym eksplanansem a eksplanandum, czy to pierwsze rzeczywiście wyjaśnia to drugie. Podstawowe czy fundamentalne w wyjaśnianiu przyczynowym są relacje zależności fizycznej, wyrażane przez odpowiednie zdania kontrfaktyczne o tym, co zdarzyłoby się przy interwencjach. Relacje wyprowadzania nie odgrywają w wyjaśnieniach roli, która byłaby niezależna czy pierwotna w stosunku do takich relacji zależności, lecz jedynie są ważne o tyle, o ile (i do tego stopnia) poprawnie reprezentują takie zależności. (Woodward 2003, 202)

Intuicyjnie można by powiedzieć, że to, jakie są te zależności, o których mowa w powyżej przytoczonym cytacie, samo będzie uzależnione od tego, czy dostarczają one zadowalających wyjaśnień zjawisk znanych z obserwacji. W ten sposób jednak otrzymalibyśmy kołowość w eksplikowaniu kategorii wyjaśniania.

Przykładem, który ma zilustrować różnicę między relacjami wynikania logicznego śledzącymi zależności fizyczne lub nie, jest przykład z kołkiem. W przypadku, gdy wyprowadzamy dedukcyjnie długość jego cienia z długości kijka oraz kąta, pod jakim pada światło, mamy – zdaniem Woodwarda – do czynienia z wynikaniem logicznym, które śledzi zależność fizyczną. W przypadku, gdy wyprowadzamy długość kijka na podstawie długości cienia i kąta światła, mamy relację wynikania logicznego, która nie śledzi zależności fizycznych.

W żaden sposób jednak nie charakteryzuje to zależności wynikania logicznego, a jedynie określa sposób dobierania przesłanek i wniosku, które – zdaniem Woodwarda – powinny być podane w układzie odpowiadającym kierunkowi zależności przyczynowej. Stwierdzamy w ten sposób ponownie brak charakterystyki zależności konceptualnej, jaka ma zachodzić w koncepcji manipulacyjnej między eksplanansem a eksplanandum. Można nawet podejrzewać, że Woodward proponuje przeciwieństwo rozwiązania przyjętego przez Hempla, w którym koncepcja przyczynowa wyjaśniania miała być szczególnym przypadkiem wyjaśniania subsumpcyjnego. Charakterystyka zależności „wyprowadzeniowej” między eksplanansem a eksplanandum sugeruje, że powielają one zależności przyczynowe, a więc, że koncepcja subsumpcyjna miałaby być szczególną postacią koncepcji wyjaśniania przyczynowego.

Po analizie roli praw oraz związku wynikania logicznego w koncepcji manipulacyjnej, przechodzę obecnie do przeanalizowania, w jaki sposób pomocna w eksplikacji wyjaśniania jest kategoria interwencji. W tej dyskusji powrócimy do krytycznych uwag dotyczących wprowadzonej przez Woodwarda kontrfaktycznego pojęcia interwencji, jakie poczyniłem w Punkcie 7.3, analizując jego koncepcję praw przyrody.

Wydaje się, że kluczowe dla proponowanego przez Woodwarda koncepcji wyjaśniania manipulacyjnego jest pojęcie interwencji. W tej koncepcji wyjaśnienie sprowadza się do podania informacji o tym, jak zmieniłoby się eksplanandum, gdyby zaistniała w danych okolicznościach wyidealizowana manipulacja eksperymentalna na wyjaśniane zjawisko. Oprócz problemów, które wskazałem w Punkcie 7.3 w odniesieniu do tak wyidealizowanego pojęcia interwencji, w poniższej dyskusji zwrócę uwagę na problematyczność odnoszenia takiego pojęcia do badań nieeksperymentalnych oraz trudności w adekwatnym ujęciu projekcyjności uogólnień przyczynowych.

9.2 Projekcyjność w procesualnej koncepcji wyjaśniania

Zastanawiając się nad stosownością aplikowania pojęcia przyczynowości rozumianego w kategoriach wyidealizowanej eksperymentalnej manipulacji w badaniach socjologicznych, Goldthorpe słusznie zwraca uwagę na fakt, że wiele z tych badań nie ma charakteru eksperymentalnego, lecz obserwacyjny. Stąd ideał postępowania naukowego:

zwolennicy zasady „nie ma przyczynowości bez manipulacji” (ang. no causation without manipulation) zalecaliby więc, aby w swoich badaniach empirycznych socjologowie tak dalece, jak tylko to możliwe, naśladowali plany eksperymentalne [...]. (Goldthorpe 2001, 5)

Trudność ta nie dotyczy wyłącznie socjologii, lecz można ją uogólnić na wiele badań nieeksperymentalnych. Warto więc bliżej prześledzić argumentację Goldthorpe’a. Rozwijając postulat idealnego postępowania dla badań nieeksperymentalnych, Goldthorpe dodaje, że naśladowanie miałyby polegać na przypisaniu nadrzędności problemowi przypisania (faktycznego lub przypuszczalnego) jednostek do grup: kontrolnej i oddziaływania, co miałyby być aproksymacją zrandomizowanego planu eksperymentalnego. Badacze powinni zwłaszcza podjąć próbę określenia, a następnie reprezentowania w swoich analizach danych za pomocą współzmiennych wszystkich wpływów na zmienną odpowiedzi, którym może podlegać w tym procesie. Zatem w badaniu wpływu wykształcenia na dochody byłoby konieczne uwzględnienie wszystkich możliwych obciążeń występujących przy przypisywaniu indywidualów do grupy oddziaływania oraz przypadkowo ustanowionych związków z rynkiem pracy w czasie prowadzenia badań.

Zasadnicza związana z tym trudność, co podkreśla Goldthorpe, dotyczy tego, skąd można wiedzieć, że zbiór współzmiennych, który zostanie w ten sposób ustalony, rzeczywiście gwarantuje, że przypisanie do grupy oddziaływania oraz odpowiedź jednostki rzeczywiście są od siebie niezależne warunkowo względem tego zbioru. Czy rzeczywiście – pyta Goldthorpe – uwzględniono, i należycie zmierzono i skontrolowano wszystkie istotne wpływy? Mimo całego arsenału środków statystycznych, który ma służyć rozwiązaniu tego problemu (Rosenbaum 2002),

trudno uniknąć wniosku – jak podkreśla Goldthorpe – że w nieeksperymentalnych badaniach prowadzonych w naukach społecznych próby określenia skutków przyczyn doprowadzą nie do rezultatów, które „nigdy nie umrą”, lecz jedynie do takich, które mają różny stopień wiarygodności. A ponieważ ta wiarygodność będzie częściowo zależna od obecnej wiedzy i teorii dotyczących danej dziedziny, które przypuszczalnie kierują wyborem współzmiennych, rezultaty te więc muszą być tymczasowe w taki sam sposób i z takich samych powodów, jak w przypadku prób określenia przyczyn skutków poprzez „analizę korelacji cząstkowych”. (2001, 6)

Zgodnie więc z tą tezą, podejście do przyczynowości w kategoriach wyidealizowanej manipulowalności w zasadzie podziela te same trudności, o których była mowa w Punkcie 7.3 w odniesieniu do modelowania przyczynowości jako stabilnej zależności przyczynowej.

Goldthorpe zwraca uwagę na dodatkową trudność, dotyczącą już tylko podejścia w kategoriach manipulowalności. Do grupy zmiennych, które nie podlegają manipulacji, należą zmienne reprezentujące cechy wchodzące w skład wsobnej, istotnej charakterystyki jednostek. W przypadku takich zmiennych, na co zwraca uwagę Goldthorpe, zastanawianie się nad przyjęciem przez nie innych wartości niż faktyczne nie ma sensu, gdyż wiązałoby się to ze zniszczeniem jednostki badania. Nie można więc, przy takim rozumieniu przyczynowości, pytać o wpływ przyczynowy rasy czy płci na osiągnięcia w wykształceniu (2001, 6). Ten właśnie problem roli przyczynowej wsobnych atrybutów Goldthorpe uznaje za zasadniczy dla podejścia opatrzonego sloganem „nie ma przyczynowości bez manipulacji”.⁷

Ponadto to podejście podważa, zdaniem Goldthorpe’a, istotność przyczynową działania. Za Hollandem (1986b) rozważmy przykład studentki, która dobrze zdała egzamin. Holland uwzględnia trzy możliwe wyjaśnienia, dlaczego uzyskała dobry wynik: (a) bo jest kobietą; (b) bo się przygotowała; (c) bo brała korepetycje u egzaminatora. Ponieważ w przypadku (c) występuje interwencja, więc sens zależności przyczynowej w podejściu manipulacyjnym jest klarowny. W przypadku (a) mamy do czynienia z wsobnym atrybutem i dlatego Holland odrzuca tu zależność przyczynową. Berk (1988, 167) podkreśla, że nie chodzi tu o płeć w sensie biologicznym, lecz jako pewien konstrukt społeczny nadbudowany na atrybucie biologicznym. Otwiera to możliwość przyczynowej interpretacji przypadku (a), gdyż płeć jako konstrukt społeczny – oraz jej społeczne postrzeganie – jest czymś w zasadzie podlegającym zmianie i manipulacji.

Za najtrudniejszy przypadek do interpretacji przy podejściu manipulacyjnym Goldthorpe uznaje (b). Intuicyjne wyjaśnienie można sformułować w ten sposób, że ta studentka planowała zdać egzamin i była przekonana, że najlepszym sposobem osiągnięcia tego celu jest przygotowanie się. Dlatego też zdecydowała się przygotować do egzaminu, a ponieważ jej przekonanie było słuszne, więc uzyskała dobry wynik. Mimo ogólności tej narracji, Goldthorpe przypisuje jej charakter przyczynowy. Holland (1986b, 955) stwierdza, że nie można przyczynowej interpretacji (b) pogodzić z podejściem manipulacyjnym, przede wszystkim ze względu na „wolitywny aspekt rzekomej przyczyny”. Goldthorpe dostrzega tu dylemat między ograniczeniem stosowalności pojęcia przyczynowości a koniecznością zreformowania dotychczasowego sposobu posługiwania się kategorią przyczynowości przez

⁷ Potwierdzenia tego wniosku Goldthorpe upatruje w pracach wielu innych autorów (m.in. Glymour 1986; Granger 1986; Berk 1988).

socjologów (2001, 7). Przyjęte przez Hollanda rozwiązanie to sceptycyzm co do teorii wnioskowań przyczynowych.

Ostateczną podstawą problemu z pogodzeniem ludzkiego działania i podejścia manipulacyjnego jest, zdaniem Goldthorpe'a, samo pojęcie interwencji. Przy podejściu manipulacyjnym interwencja jest traktowana tak, jak w przypadku systemów studiowanych w naukach przyrodniczych. Jedyne miejsce, gdzie dopuszcza się ludzkie („celowe”) działanie, jest rola eksperymentatora – gdy zrealizowano już interwencję, cały system ma „zachowywać się po prostu jak bakteria reagująca na środek chemiczny” (2001, 8). Wszelkim innym działaniom ludzkim, a w szczególności podejmowanym w codziennej realizacji zamierzeń przez indywiduala, odmawia się istotności przyczynowej, co można uznać za swego rodzaju paradoks podejścia akcentującego właśnie manipulowalność.⁸

Dwie skrajne reakcje na ten problem reprezentują prace Sobela (1995, 1996) oraz Lieberzona (1985). Sobel akceptuje podejście manipulacyjne jako wyznaczające standardy wnioskowań przyczynowych. W konsekwencji badania powinny przebiegać wedle planu eksperymentalnego lub quasi-eksperymentalnego. Jeśli to nie jest możliwe, to badania nieeksperymentalne powinny zaakceptować ten właśnie schemat pojęciowy dla oceny prawomocności wnioskowań przyczynowych, mimo ograniczeń, jakie się z tym wiążą.

Lieberzon natomiast zasadniczo odrzuca jakąkolwiek próbę narzucenia modelu eksperymentalnego badaniom nieeksperymentalnym, twierdząc, że jest to przejaw scjentyzmu i przeceniania formy, a nie substancji metody naukowej. W konsekwencji badacze muszą wypracować własny sposób rozstrzygnięcia problemów dotyczących zależności przyczynowych, który będzie stosowny do prowadzonych badań i rzeczywiście podejmowanych problemów.

Inne stanowisko reprezentują metodolodzy z Nuffield College w Oksfordzie, a więc Cox, Goldthorpe i Wermuth. Zasadniczo przyjmują oni statystyczne kryteria jako warunek konieczny zależności przyczynowej, lecz niewystarczający. Cox i Wermuth (1996, 220-21) postulują jawne wprowadzenie „pojęcia tkwiącego u podstaw [tej zależności] procesu”, które

⁸ Podobny problem Goldthorpe dostrzega w dyskusjach filozoficznych, zainspirowanych przez tezę neowittgensteinistów, że przyczyny i skutki muszą być logicznie niezależne, co wykluczałoby raczej podjęcia działania z zakresu możliwych przyczyn (2001, 16 p. 6). Polemikę z takim poglądem podjęli Toulmin (1970), Mackie (1974, R. 11) i Davidson (1980, R. 1 i 14).

dotyczy „poziomu obserwacyjnego leżącego głębiej niż ten związany z danymi, które są analizowane”. Podobną uwagę można znaleźć w pracy Simona i Iwasaki (1988, 157), którzy podkreślają, że asocjacja statystyczna powstaje dzięki pewnemu „mechanizmowi”, działającemu na poziomie „mikroskopowym” w stosunku do tego poziomu, na którym stwierdza się istnienie takiej korelacji.⁹ Innymi słowy, proces, który istnieje w czasie i przestrzeni, choć nie musi być obserwowalny, faktycznie generuje efekt przyczynowy jednej zmiennej na drugą, czego manifestacją jest zależność statystyczna stwierdzona empirycznie w danych. Simon i Iwasaki zgadzają się jednak, że takie procesy przyczynowe, które wyjaśniają miejsca niedookreślone przez analizę statystyczną, nigdy nie są stwierdzane definitywnie – muszą być bowiem podatne na testowanie empiryczne i nawet jeśli są w ten sposób potwierdzone, to i tak trzeba liczyć się z możliwością podania wyjaśnień w bardziej subtelnym kategoriach, na głębszym poziomie.

Jako wzorcowy przykład takiego właśnie podejścia Goldthorpe (2001, 9) podaje epidemiologię (Hill 1937/1991, 1965) i badanie zależności między paleniem papierosów a rakiem płuc. Analizy statystyczne danych obserwacyjnych ujawniły mocną korelację między paleniem a rakiem płuc oraz jej stabilność – uwzględnienie szeregu możliwych wspólnych przyczyn nie zmieniło tej zależności statystycznej. Jednak, jak podkreśla Goldthorpe, kluczowe dla stwierdzenia przyczynowego charakteru tej zależności było podanie procesu generującego tę zależność statystyczną przez wyizolowanie substancji rakotwórczych w dymie papierosowym, histopatologicznych dowodów z nabłonka oskrzeli itp. W tym kontekście Freedman (1997, 129) zwraca uwagę na zróżnicowanie źródeł danych potwierdzających zaproponowany proces generujący zależność statystyczną oraz wieloaspektową zgodność różnych studiów z danymi statystycznymi.¹⁰

⁹ Por. też (Freedman 1991, 1992a, 1992b).

¹⁰ Wydaje się jednak, że Freedman nie wprowadza supozycji ontologicznej, a jedynie ogranicza się do stwierdzenia relacji między danymi a postawioną hipotezą. Podobny wniosek wyciąga Cox, który przytacza także listę B. Hilla warunków, mających zagwarantować poprawność wnioskowań przyczynowych w badaniach nieeksperymentalnych proceduralnie, niemechanicznie (2001, 70). Goldthorpe natomiast (2001, 10n) uogólnia i podaje trzy kroki w określaniu zależności przyczynowej: określenie zjawisk, które stanowią eksplanandum, hipoteza o procesie na poziomie działania społecznego oraz testowanie hipotezy. Te kroki występują w każdym podejściu do modelowania przyczynowości, natomiast zasadnicza różnica to „proces”, który oznacza tu „model przyczynowy”. Potwierdza ten wniosek także analiza podejścia procesualnego jako sprowadzalnego do dwóch pozostałych.

Przywołując opinię Freedmana, Goldthorpe zdaje się jednak łączyć ze sobą dwie różne tezy: jedna dotyczy założenia o istnieniu procesu przyczynowego, którego manifestacją są zależności statystyczne, druga teza natomiast głosi konieczność odwołania się wiedzy teoretycznej dotyczącej danej dziedziny przy orzekaniu zależności przyczynowej i niewystarczalność samych analiz statystycznych. Teza druga jest zgodna z manipulacyjnym podejściem do przyczynowości, natomiast pierwsza – jak podkreśla to sam Goldthorpe – w pewnym zakresie jest niezgodna.

Wróćmy do przykładu studentki. Goldthorpe stwierdza, że przy przyjęciu założenia o procesie przyczynowym, zarówno (a), jak i (b), mają interpretację przyczynową, z którą – jak zaobserwowaliśmy wcześniej – są zasadnicze trudności przy podejściu manipulacyjnym.¹¹ Przy podejściu procesualnym płeć studentki (w sensie biologicznym) może pełnić rolę przyczynową, gdyż można podać wyjaśnienie odwołujące się do faktu, że mózgi kobiet ewoluowały w taki sposób, który daje im przewagę nad mężczyznami w takim rodzaju zadań, które należało rozwiązać na egzaminie (Goldthorpe 2001, 9).

W kontekście nauk społecznych natomiast przyczynowa istotność działania nie tylko nie jest problematyczna, ale jest „standardowym sposobem konstruowania wyjaśnienia procesów przyczynowych” (Goldthorpe 2001, 9). I tak, w przypadku (b) podanie procesu, który prowadzi od zamierzeń studentki i jej przekonań poprzez przygotowanie się do egzaminu i dobry wynik, stanowi wyjaśnienie przyczynowe tego ostatniego.

W stanowisku Goldthorpe’a wyróżnione zostały wyżej dwie tezy: jedna dotycząca odwoływania się do wiedzy teoretycznej, wykraczającej poza dane statystyczne, a druga dotycząca istnienia procesu przyczynowego. Przestrzegając tego odróżnienia, w analizie przykładu studentki zajmę się tylko drugą.

W przypadku (a) odpowiedź naszkicowana przez Goldthorpe’a na gruncie podejścia procesualnego jest w zasadzie zastosowaniem techniki zmiany tematu. Po pierwsze, nie jest jasne, dlaczego w podanym wyjaśnieniu jest mowa o przewadze kobiet nad mężczyznami, skoro – jak się wydaje – pierwotne sformułowanie przykładu nie jest w nieuchronny sposób z tym związane. Po drugie, zwróćmy uwagę, że pierwotnie kontekst oczekiwanego wyjaśnienia dotyczył badań społecznych, nieeksperymentalnych. W naszkicowanym przez Goldthorpe’a procesie mowa jest jednak w zasadzie o płasz-

¹¹ Podana przez Goldthorpe’a dystynkcja procesualne-manipulacyjne jest niejasna, wzięwszy pod uwagę, że sztandarowe podejście manipulacyjne w postaci sieci bayesowskich jest określane także jako podawanie „modelu procesu generującego dane” (Pearl 2000, 140).

czyźnie neuronalnej i kognitywnej, co więcej – w kontekście, który jak najbardziej dopuszcza badania eksperymentalne. W ten sposób zmieniony przykład nie nastęrcza jednak trudności przy podejściu manipulacyjnym.

Nieco podobnie można przeanalizować przykład (b). Mamy tu nie tylko pewne wyjaśnienie przyczynowe w kategoriach zmiennych obserwowalnych (ewentualnie z udziałem zmiennych ukrytych), lecz także hipotezę ontologiczną – choć innego rodzaju. Jak w przypadku (a) Goldthorpe postulował przyjęcie pewnego procesu na poziomie neuronalnym, tak w przypadku (b) odwołuje się do istnienia nieobserwowalnych kategorii, jak zamierzenia, wolitywność czy intencjonalność działań studentki. Podane przez Goldthorpe'a wyjaśnienie w tych kategoriach można zinterpretować na dwa sposoby. W pierwszym przypadku takie wyjaśnienie potraktowane może być jak pewien model zależności przyczynowych, w drugim – jako odwołanie się w interpretacji danego modelu do wiedzy teoretycznej. Pierwszy przypadek zostaje poza kryterium identyfikowalności, co w zasadzie wyklucza go z metodologii badań nieeksperymentalnych, gdyż nie pozwalałby określić efektu przyczynowego, jaki miało przygotowanie się do egzaminu na dobry wynik studentki.

Drugi przypadek nie ujawnia nic specyficznego dla podejścia procesualnego, gdyż to samo wyjaśnienie można podać przy podejściu manipulacyjnym, przyjmując jednocześnie zastrzeżenia Freedmana, który postuluje odwołanie się do wiedzy teoretycznej ze względu na niewystarczalność analiz statystycznych przy określaniu zależności przyczynowych.

Podsumowując tę dyskusję, można stwierdzić, że podejście procesualne nie stanowi zasadniczo odrębnego podejścia do przyczynowości poza podejściem manipulacyjnym oraz poza stanowiskiem Freedmana, który postuluje odwoływanie się do wiedzy teoretycznej o badanej dziedzinie. Propozycje zgłaszane przez Goldthorpe'a na gruncie podejścia procesualnego bądź redukują się do podania innego – niż wcześniej zaproponowany – modelu przyczynowego na gruncie podejścia manipulacyjnego, bądź do podkreślenia – za Freedmanem – roli wiedzy teoretycznej jako jednego z warunków koniecznych stwierdzenia zachodzenia zależności przyczynowych.

Obawy zgłaszane przez Goldthorpe'a i Liebersona mają jednak swój odpowiednik w pewnym problemie koncepcji manipulacyjnej wyjaśniania, jaki można podnieść z perspektywy proceduralnej. Zgodnie z tą koncepcją interwencja (w obrębie pewnej dopuszczalnej klasy) na badaną rzeczywistość może prowadzić do określonych zmian eksplanandum. Jednocześnie zależność przyczynowa, która ma być podstawą podanego wyjaśnienia, po-

winna charakteryzować się niezmienniczością na efekty takich interwencji. Interwencje w obszarze dopuszczalnej klasy wyznaczają więc, jakie zmiany byłyby możliwe przy zachowaniu niezmienniczości obdarzonej mocą eksplanacyjną zależności przyczynowej oraz jaka jest przestrzeń możliwych alternatyw dla podanego eksplanandum.

W stosunku do takiego obrazu wyjaśniania obawy Goldthorpe'a i Lieberzona można wyrazić jako zastrzeżenie co do tej przestrzeni alternatyw. Jeśli bowiem jako możliwa zdefiniowana zostanie interwencja dopuszczalna kontrfaktycznie, to rzeczywiście stajemy wobec problemu, że w koncepcji manipulacyjnej nie jest przestrzegane odróżnienie kontekstu badań eksperymentalnych i nieeksperymentalnych. Jedne i drugie bowiem są zrównane w sposobie wyznaczania przestrzeni możliwych alternatyw nie przez *faktycznie* realizowalne interwencje, lecz dopuszczalne kontrfaktycznie, a więc konceptualnie.

Takie jednak rozumienie interwencji, oprócz problemów wskazanych w Punkcie 7.3, otwiera także drogę do sformułowanego radykalnie przez Goodmana problemu przepełnienia zbioru możliwych alternatyw. W tym zbiorze znajdują się bowiem zarówno możliwości projekcyjne, jak i nieprojekcyjne. To ostatecznie prowadzić będzie do trywializacji pojęcia wyjaśniania.

Problemu przepełnienia i trywializacji unika się, jak wynika z dyskusji prowadzonej w kolejnym punkcie, przyjmując proceduralną charakterystykę interwencji (zbierania danych w obszarze pola interwencji danej zależności przyczynowej), która prowadzi do wyodrębnienia uogólnień projekcyjnych.

9.3 Wyjaśnianie przyczynowe jako subsumpcja proceduralna

Wyjaśnianie przesądza o dynamice badań naukowych. Konkluduje zdobytą dotąd wiedzę, a zarazem ujawnia nowe pytania i kierunki poszukiwań. Jak pokażę niżej, źródłem tej dynamiki wyjaśniania jest integracja proceduralności modeli przyczynowych z projekcyjnością uogólnień. Od modeli przyczynowych żąda się, by spełniały kryterium proceduralne zastosowane do danych interwencyjnych, pochodzących z interakcji z badaną rzeczywistością. Od uogólnień projekcyjnych natomiast żąda się, by poza obszar dotychczas objęty interakcją badawczą przenosiły znane zależności. Takie uogólnienia mają z jednej strony podstawę w proceduralnym modelu przyczynowym, z drugiej natomiast wykraczają poza dziedzinę, dla której ten model został określony. Wyjaśnienie czerpie więc swoją dynamikę z połączenia dwóch odmiennych tendencji ludzkiego poznania: zaliczenia danego

zjawiska do prawidłowości, znanej z modelu przyczynowego, ale jednocześnie takiej, która jest projekcyjna i dotyczy szerszej klasy zjawisk niż te, z którymi dotychczas podjęto interakcję badawczą.

Ten ogólny zarys proceduralnej koncepcji wyjaśniania sugeruje sposób, w jaki łączą się w niej wyjaśnianie subsumpcyjne z przyczynowym. Między eksplanandum a eksplanansem zachodzą bowiem dwa rodzaje związków: związek subsumpcji, czyli podporządkowania pewnej prawidłowości, oraz związek asymetrii, korespondujący z asymetrią zależności przyczynowej między zjawiskami, do których odnosi się eksplanandum i eksplanans. W dalszej części tego punktu staram się wykazać, że koncepcja proceduralna wyjaśniania pozwala adekwatnie połączyć oba rodzaje związków przy jednoczesnym zachowaniu ich odrębności.

Trudność z podaniem eksplikacji zależności subsumpcji i zależności przyczynowej, które pozwalałyby na zintegrowanie tych dwóch zależności w wyjaśnianiu, a jednocześnie zachowywałyby ich odrębność, prześledziłem we wcześniejszych partiach tekstu w odniesieniu do wielu współczesnych koncepcji wyjaśniania. Koncepcja DN Hempla, a za nią także unifikacjonistyczna Kitchera, oparta była na tezie redukcyjnej, że zależność przyczynowa jest szczególnym przypadkiem zależności subsumpcyjnej. Ponadto przyjęta przez Hempla redukcja prowadziła do konsekwencji, które określiłem jako totalny deskrypcjonizm (wymóg wyrażenia słownego wszystkich okoliczności istotnych dla wyjaśnianego zdarzenia) oraz determinizm przyczynowy (pełne podanie przyczyn determinuje zaistnienie wyjaśnianego zdarzenia, podobnie jak eksplanandum wynika logicznie z zawierającego totalny opis eksplanansa). Koncepcja manipulacyjna Woodwarda przeciwnie – traktowała zależność subsumpcji jako szczególny przypadek zależności przyczynowej.

Wśród dotychczas proponowanych koncepcji wyjaśniania znalazła się także inna trudność, której nie sposób wskazać, nie przyjmując – przynajmniej w podstawowym wymiarze – proponowanego w tej pracy podejścia proceduralnego. Chodzi mianowicie o odróżnione we Wprowadzeniu dwa typy analiz zależności przyczynowych. W dyskusji tej zależności należy oddzielić zadanie określenia kryteriów determinujących, kiedy mamy do czynienia z zależnością przyczynową, od filozoficznej zadania – eksplikacji natury przyczynowości. W tej pracy podejmuję tylko pierwsze zadanie, a zasadniczą rolę w jego realizacji pełni zaproponowane w Punkcie 2.3 kryterium proceduralne. Jak pokazałem, pozostaje ono w zgodzie z szeroką gamą

rozstrzygnięć filozoficznych w eksplikowaniu znaczenia zależności przyczynowej.

Przykładem koncepcji wyjaśniania, w której wyraźnie można zaobserwować problem pomieszania podejścia kryterialnego z filozoficznym, jest Salmona koncepcja wyjaśniania CM. Obserwujemy w niej napięcie, by z jednej strony – określić warunki relewancji przyczynowej eksplanansu w stosunku do eksplanandum i uniknąć związanych z tym trudności wcześniejszej koncepcji IS Hempla. Z drugiej strony, Salmon stara się zastosować neo-Hume'owską analizę kategorii przyczynowości, redukując ją do kategorii następstwa czasowego i regularności. Pomieszenie tych dwóch zadań prowadzi do koncepcji, która – jak pokazałem w poprzednich punktach – nie realizuje żadnego z tych dwóch zadań zadowalająco.

Przejdę obecnie do szczegółowej dyskusji proceduralnej koncepcji wyjaśniania, która – jak sądzę – unika powyższych trudności. Rozpocznę od przywołania przykładu badań Snowa nad przyczyną epidemii cholery w Londynie w połowie XIX wieku. Czy zaproponowany przez Snowa mechanizm rozwoju i rozprzestrzeniania się tej choroby można uznać za wyjaśnienie?

Jak stwierdził to jego bliski przyjaciel Richardson, Snow już podczas pierwszej fali epidemii w latach 1830. w Anglii na podstawie symptomów choroby doszedł do wniosku, że powoduje ją żywy pasożytniczy organizm, który przenosi się wraz z wodą lub żywnością, atakuje bezpośrednio układ pokarmowy, a obserwowane zmiany w krwi są tego następstwem, które bezpośrednio prowadzi do dysfunkcji różnych organów i śmierci.¹² Podkreślmy, że Snow nie dysponował ani obserwacjami owego organizmu, którego istnienie postulował, ani odpowiednimi danymi z populacji. Nie można więc w tym przypadku mówić, że zaistniała taka interakcja z badaną rzeczywistością, która pozwoliłaby określić, czy ten model przyczynowy spełnia kryterium proceduralne. Uogólnienie, które na tym etapie badań zaproponował Snow, próbując wyjaśnić przyczynę cholery, nie było więc projekcyjne i nie stanowiło zadowalającego wyjaśnienia. Nawet jeśli wskazany przez Snowa model przyczynowy okazałby się – za sprawą badań Kocha i jego następców – słuszny, to w latach 1830., a więc wtedy, kiedy Snow zaproponował go po raz pierwszy, nie był to model, który – zgodnie z kryterium proceduralnym – uznalibyśmy za przyczynowy, gdyż nie był ustalony na podstawie odpowiedniej interakcji z badaną rzeczywistością i nie spełniał kryterium proce-

¹² Por. (Sack i in. 2004, 224-27).

duralnego. Mówiąc inaczej, mimo że związek subsumpcji, jaki wskazał w tym przypadku Snow, między przypadkami zachorowań a mechanizmem rozwoju i przenoszenia choroby, okazał się słuszny, to jednak nie zostało wykazane zachodzenie zależności przyczynowej między zjawiskiem wyjaśnianym a wyjaśniającym. To właśnie było motywacją podjęcia przez Snowa wielu wysiłków, zmierzających do wykazania, że postulowany przez niego model zależności spełnia jednak kryterium proceduralne.

W niektórych przypadkach, jak choćby w odniesieniu do rodziny, która wspólnie zjadła mięso skażone cholera, Snow podawał szereg informacji kontrfaktycznych, określających jakościowo efekty manipulacji: co byłoby, gdyby nie zjedli tego mięsa lub gdyby zjedli je po odpowiednim przygotowaniu, np. ugotowaniu. Można by więc zastanawiać się, czy na gruncie manipulacyjnej koncepcji wyjaśniania z uwagi na te informacje kontrfaktyczne, opisujące efekty możliwych interwencji, nie mamy do czynienia z wyjaśnieniem. Odpowiedź jest jednak negatywna. W tym przypadku Snow *implicite* przyjmował uogólnienie, że przygotowanie pokarmu, które nie usuwa bakterii cholery oraz obecność tych bakterii w zakażonym mięsie, były przyczyną zachorowania. Przy braku modelu przyczynowego można byłoby mówić co najwyżej o trafnym wskazaniu związku subsumpcji, który przy braku danych pozwalających stwierdzić zależność przyczynową nie jest jednak wystarczającą podstawą wyjaśnienia.

Zdjęcie ramienia z pompy przy Broad Street na polecenie Snowa dostarczyło danych zgodnych z postulowanym przez Snowa modelem przyczyn cholery. Ze względu jednak na emigrację niekontrolowanej liczby osób z tego obszaru nie są to dane wystarczające do spełnienia kryterium proceduralnego dla tego modelu. Podobnie zresztą dla wielu drobiazgowych przypadków, które omawia Snow.

Dane z pola interwencyjnego, które były wystarczające do spełnienia kryterium proceduralnego dla modelu przyczyn cholery, Snow zdobył, analizując dane o dostawcach wody w południowych dzielnicach Londynu w czasie epidemii w latach 1854-55. Jak podkreśla, był to eksperyment naturalny, który objął populację blisko trzystu tysięcy osób.

Czy zatem dysponując takim modelem przyczynowym oraz zależnością subsumpcyjną, których słuszność Snow starał się wykazać w stosunku do wszystkich znanych mu przypadków zachorowań na cholera, można uznać, że zostało podane wyjaśnienie przyczyn tej choroby? Czy następujące wnioskowanie Snowa może być uznane za wyjaśnienie: Wprowadzanie do układu

pokarmowego przez człowieka niewidzialnego organizmu prowadzi do zakażenia cholera, a ta w wielu przypadkach kończy się śmiercią.

Zacznę od wymogów, jakie stawia koncepcja DN, a którym proponowane przez Snowa wyjaśnienie nie może sprostać. W podanym wyjaśnieniu Snow nie przywołuje praw naukowych, a jedynie określone prawidłowości. Zwolennicy alternatywnego wyjaśnienia przyczyn cholery uznaliby prawidłowość zachodzącą między piciem skażonej wody a zachorowalnością na cholera: Wśród pijących skażoną wodę kilkakrotnie nasila się liczba przypadków zachorowań na tę chorobę. Takie uogólnienie nie jest oczywiście bezwyjątkowym prawem, jakiego wymaga wyjaśnienie DN.

Inna trudność związana jest z zachodzeniem związku wynikania logicznego między eksplanansem a eksplanandum. Dla każdego znanego sobie przypadku Snow był w stanie, posługując się wprowadzonym przez siebie modelem zależności przyczynowych, podać wystarczający opis zaistniałych warunków, na podstawie których można było podporządkować dany przypadek tej prawidłowości. Trudno jednak byłoby sprowadzić rozważania Snowa wyłącznie do wyróżnienia czterech klas w partycji badanych zjawisk (uwzględniając spożycie skażonej wody i zachorowanie na cholera) i klasyfikowanie przypadków do jednej z tych klas partycji. Nie można więc w tym przypadku mówić o wynikaniu logicznym, a tym bardziej, że spełnione jest wprowadzone przez Hempla założenie totalnego deskrypcjonizmu.

Nawet domniemana przez Hempla w koncepcji DS reguła uznawania zdań o wysokim prawdopodobieństwie nie ma tutaj zastosowania do obrony tezy, że mamy do czynienia z „wynikaniem indukcyjnym”. Weźmy pod uwagę przykład, gdy Snow wyjaśniał przypadek zachorowania na cholera z powodu tego, że jeden z członków rodziny dotykał ubrań zmarłej na cholera siostry i przeniósł te bakterie na spożywane w rodzinie pożywienie. Uogólnienie „Każdy, kto spożywa jedzenie z rękoma skażonymi bakterią cholera, zachoruje na tę chorobę” nie gwarantuje nawet wysokiego prawdopodobieństwa wniosku, że taka to a taka osoba, która jadła posiłek rękoma skażonymi tą bakterią, zachoruje na cholera.

Z punktu widzenia koncepcji DN można podnieść jeszcze inny problem w stosunku do proponowanego przez Snowa wyjaśnienia przyczyny cholera. Mimo że organizm, o którym Snow pisał jako przyczynie cholera, znany był z obserwacji, to jednak brak było dostatecznego dowodu, że to właśnie ten opisany i uwieczniony na jednym z rysunków w raporcie komisji CSI mi-

kroorganizm¹³ jest opisywanym przez Snowa pasożytem. Można więc twierdzić, że zaproponowany przez Snowa mechanizm biologiczny rozwoju cholery w tym sensie nie spełniał wymogu testowalności empirycznej, jaki związany jest z koncepcją DN.

Hempel wprowadził warunek testowalności empirycznej eksplanansa jako warunek syntaktyczny. W związku z tym nie wydaje się, by z tego warunku wynikało, że konieczne jest podanie aktualnie dostępnych danych, aby daną hipotezę uznać za testowalną. W tym więc słabszym – czy syntaktycznym – sensie testowalność empiryczna postulowanego przez Snowa mechanizmu rozwoju cholery została potwierdzona przez prowadzone trzydzieści lat później badania Kocha. Zaobserwował on ten mikroorganizm, wyizolował go i rozmnożył w czystej kulturze, a następnie wprowadził do organizmu i ponownie wyizolował, aby pokazać, że spełnione są postulaty¹⁴ wystarczające do stwierdzenia, że to właśnie *vibrio cholerae* jest przyczyną cholery. Ponadto Koch na podstawie dokonywanych w czasie sekcji zwłok chorych zmarłych na cholere obserwacji potwierdził postulowany przez Snowa wpływ tej bakterii na zmiany w składzie chemicznym krwi i uznał odwodnienie za bezpośrednią przyczynę zgonu.

Mimo że podany przez Snowa eksplanans spełnia niektóre z wymogów koncepcji DN, jak testowalność empiryczna czy prawdziwość, to jednak nie jest przykładem wyjaśnienia w sensie koncepcji DN czy DS. Nie występują w nim bowiem ani prawa przyrody, ani związek wynikania logicznego między eksplanansem a eksplanandum. Z uwagi jednak na trudności koncepcji DN i jej pochodnych, na jakie zwróciłem uwagę w Punkcie 8.1, można zastanawiać się, czy uznać je za trudności wyjaśnienia zaproponowanego przez Snowa, czy też raczej za powód nieadekwatności koncepcji DN i jej własnych problemów.

Powszechnie przykład badań Snowa kwalifikowany jest w statystyce (Freedman 1999) i epidemiologii (Susser 1973, 54-56; Turner 1997, 36-40) jako wyjaśnienie przyczyn epidemii cholery. Jeśli tak, znaczyłoby to, że mamy tu do czynienia z wyjaśnieniem w innym sensie niż eksplikowany przez koncepcję DN.

Koncepcja subsumpcji proceduralnej modyfikuje dwa zasadnicze elementy koncepcji DN: kategorię praw jako podstawowego składnika eksplanandum oraz wymóg wynikania logicznego między eksplanandum a eksplananssem. Z dyskusji prowadzonej w Rodziale 6. wynika, że tę samą, co

¹³ Por. (Paneth i in. 1998).

¹⁴ Por. (Zieliński i Mazurek 2002, 279).

prawa, rolę eksplanacyjną pełnią uogólnienia projekcyjne. Stąd też nie można żądać od uogólnień podawanych w wyjaśnieniach spełnienia – uznanych w poprzednich analizach za dodatkowe – warunków, jakie są wymagane do tego, by dane uogólnienie uznać za prawo naukowe.

W dotychczasowych rozważaniach zwracałem uwagę na różne próby osłabienia lub wyeliminowania związku wynikania logicznego. W Salmona koncepcji wyjaśniania CM, która miała na celu takie ulepszenie koncepcji DN i jej pochodnych, aby zachować odrębność koncepcji wyjaśniania subsumpcyjnego i przyczynowego, związek wynikania logicznego został zakwestionowany. Salmon uznał, że porównanie prawdopodobieństwa pierwotnego eksplanandum z jego prawdopodobieństwem wtórnym w klasie odniesienia wyznaczonej przez istotne cechy, nie ma w ogóle charakteru wnioskowania (Salmon 1971b). A zatem nie powstaje także problem dyskusowania zależności logicznej między eksplanandum a eksplananssem. W późniejszej jednak fazie, pod wpływem wielu krytyk tej koncepcji, Salmon uznał potrzebę wprowadzenia – przynajmniej w niektórych przypadkach – wyjaśnień o postaci wnioskowań.

Manipulacyjna koncepcja wyjaśniania Woodwarda, jak wynika z dyskusji w Rozdziale 7., w zasadzie nie podaje jasnej charakterystyki związku między eksplananssem a eksplanandum. Charakterystyka tego związku jako „śledzącego czy odzwierciedlającego” zależności przyczynowe w zasadzie – jak starałem się pokazać – prowadzi do jego zredukowania do tychże zależności. I ostatecznie, wbrew temu, co sugerowałby Woodward, związek wynikania logicznego traktowany jest tu nie jako szczególny przypadek szerszej klasy zależności subsumpcji, lecz jako jej epifenomen. Zależność subsumpcji zachodzi bowiem między eksplanandem a eksplananssem niezależnie od tego, czy zachodzi związek wynikania logicznego.

Dyskusję wyjaśniania jako subsumpcji proceduralnej rozpocznę od krótkiej prezentacji tego zamysłu. Model przyczynowy, który otrzymujemy dzięki interakcji badawczej z rzeczywistością, spełniający kryterium proceduralne, stanowi podstawę projekcji partycji badanej rzeczywistości.¹⁵ Zaliczenie – na podstawie posiadanej wiedzy teoretycznej lub ogólnej – zdarzenia opisywanego w eksplanandum do jednej z klas partycji pozwala na jego wyjaśnienie. W przypadku, gdy są wystarczające podstawy do stwierdzenia, że podana partycja jest wyczerpująca i gdyby uwzględniono wszyst-

¹⁵ Liczba klas tej partycji będzie uzależniona od uwzględnianych zjawisk V i ich rodzaju (a zwłaszcza zbioru wartości, które mogą one przyjmować).

kie relewantne zjawiska,¹⁶ wyjaśnienie jest rozumowaniem *tollendo ponens* (Czeżowski 1968, 197-200).

W ogólnym przypadku natomiast wybór klasy partycji, której podporządkowane zostanie wyjaśnianie zjawisko, polega na ustaleniu, przy której z tych klas partycji – uwzględniając dostępną wiedzę – prawdopodobieństwo wystąpienia zjawiska, opisywanego w eksplanandum, jest największe. Tego rodzaju rozumowanie znane jest w statystyce jako maksymalizacja wiarygodności.¹⁷

Dla uchwycenia specyfiki subsumpcji proceduralnej przywołam prosty¹⁸ przykład wyjaśniania DN. Alfred jest czarny, bo jest krukiem, i wszystkie kruki są czarne. To wyjaśnienie w koncepcji DN ma schematycznie postać:

$$\begin{array}{l} (1) \quad \forall_x K(x) \rightarrow C(x) \\ (2) \quad K(a) \\ \hline (3) \quad C(a) \end{array}$$

Rysunek 9.1.

Zwróćmy uwagę, że przesłanka większa (1) jest niczym innym jak wprowadzeniem partycji badanej rzeczywistości na cztery podklasy. Jeśli założyć, że przyjęta milcząco wiedza implikuje pustość *wszystkich* logicznie możliwych klas partycji wyznaczonych w badanej rzeczywistości przez relewantne własności (tu: bycie krukiem i bycie czarnym) – z wyjątkiem tej, do której odnosi się przesłanka (1) – to można uznać, że podane wyjaśnienie logicznie implikuje wniosek (3).

W przypadku, gdyby okazało się, że nie uwzględniono wszystkich relewantnych własności, można, oczywiście, przesłankę (1) uzupełnić o brakujące predykaty ze zbioru V . Wprowadzenie dodatkowych predykatów wzbogaci liczbę logicznie możliwych klas o $2^V - 4$, co nadal jest liczbą skończoną. Podobnie jednak jak wcześniej, *nie można* zrezygnować z wymogu, że niepustość klas wyróżnionych w zmodyfikowanej przesłance (1) musi implikować pustość wszystkich pozostałych logicznie możliwych klas.

¹⁶ Jak zwrócę na to uwagę niżej, takie założenie, że z projekcji partycji w oparciu o wiedzę teoretyczną lub ogólną wynika eksplanandum, jest mocnym uproszczeniem i niesie za sobą pewne nieoczekiwane konsekwencje.

¹⁷ Por. (Backe 1999).

¹⁸ Zawieszam w tym miejscu wątpliwości co do tego, czy własność bycia krukiem jest odpowiedniego rodzaju własnością do wyjaśniania koloru upierzenia ptaka.

Uwzględniając ten komentarz, przedstawię przykład w postaci, która ułatwi jego uogólnienie.¹⁹ Przyjmę następujące oznaczenia: $P_1 = [x: C(x) \wedge K(x)]$; $P_2 = [x: C(x) \wedge \neg K(x)]$; $P_3 = [x: \neg C(x) \wedge \neg K(x)]$ oraz $P_4 = [x: \neg C(x) \wedge K(x)]$.

$$(1) \quad \forall_x x \in (P_1 \vee P_2 \vee P_3)$$

$$(2) \quad a \in (P_1 \vee P_4)$$

$$(3) \quad a \in P_1$$

Rysunek 9.2.

Przesłanka (1) przedstawiona na Rysunku 9.2 określa sposób dokonania partycji w pewnym zbiorze przedmiotów P (np. ptaków), z uwzględnieniem rezultatów interakcji badawczej z przedmiotami ze zbioru P . I tak, z logicznie możliwych kombinacji została – w oparciu o rezultaty interakcji badawczej – wykluczona partycja P_4 . Przy uogólnieniu sposobu sformułowania tej przesłanki, należy wziąć pod uwagę, po pierwsze, przypadek, gdy nie wszystkie zmienne są dychotomiczne (binarne), a po drugie, przypadek, gdy zmienne przyjmują wartości rzeczywiste w określonym przedziale.

Oba rodzaje zmiennych występują w modelach przyczynowych reprezentujących zależności, zgodnie z kryterium proceduralnym. Wykorzystując terminologię wprowadzoną w Punkcie 4.4, w ustalaniu partycji danego zbioru przedmiotów Y należałoby uwzględnić te spośród analizowanych rodzajów zjawisk V , z którymi Y pozostaje w zależnościach przyczynowych. Ten podzbiór V określa się jako pokrycie Markowa dla Y w zbiorze V : $Y \perp\!\!\!\perp V \setminus Y \mid \text{bl}(Y)$. Zmienne w obszarze pokrycia Markowa $\text{bl}(Y)$ gwarantują Y niezależność probabilistyczną od pozostałych zmiennych w V . Zatem przy konstrukcji partycji Y należy uwzględnić tylko zmienne z pokrycia Markowa $\text{bl}(Y)$.²⁰

Sposób wprowadzenia partycji przez zmienne z pokrycia Markowa w zbiorze Y może być przeprowadzony z uwagi na własności brzegowego rozkładu prawdopodobieństwa: $P(Y) = \sum_i P(Y \mid X_i) P(X_i)$. Uwzględniając macierze kowariancji lub charakter funkcji rozkładu dla poszczególnych członów tego iloczynu, można wyodrębnić człony partycji w zbiorze Y .

¹⁹ Podobny sposób uogólnienia zaproponował Amsterdamski (1983, 101).

²⁰ Por. (Gordon i Vichi 1998).

Przejdźmy teraz do przesłanki (2) z Rysunku 9.2. Charakteryzuje ona pewną jednostkę w kategoriach, jakie dopuszczane są przez zmienne w V oraz zbiór ich wartości. Należy podkreślić, że ta charakterystyka jest niezależna od stwierdzonych w interakcji badawczej w oparciu o kryterium proceduralne zależności między zmiennymi w V .²¹ W szczególności ta charakterystyka jest niezależna od tego, jak zostało w modelu przyczynowym dla V określone pokrycie Markowa dla zbioru Y , do którego należy charakteryzowana jednostka.²²

W wyjaśnianiu jako subsumpcji proceduralnej korzysta się, jak podałem wyżej, z dwóch rodzajów przesłanek o odmiennym charakterze. Dlatego też można sądzić, że koncepcja ta pozwala na znalezienie drogi pośredniej między zasugerowaną przez Czeżowskiego dychotomią:

Podział logiczny jest typową dla racjonalistycznego postępowania metodą badania wnętrza zakresów metoda ta daje dodatnie rezultaty tylko tak długo, póki system podziałów nie staje się zbyt skomplikowany. Empirystyczne podejście pozwala poznać strukturę elementarną badanego zakresu rzeczywistości w sposób o wiele bardziej subtelny przez zastosowanie metody statystycznej, związanej najściślej z indukcją i pokrewnymi jej odmianami rozumowania prawdopodobieństwowego. Racjonalista układa przedmioty w klasy, empiryk potrafi uporządkować je inaczej jeszcze, ustawiając je w ciągi lub szeregi, które pozwalają *przedstawić zmienność* każdej badanej własności w obrębie zasięgu jej występowania. (1949/1958, 32; podkreślenie – P. K.).

Zmienność, dodajmy, szczególnie tu istotną, a mianowicie spowodowaną podjęciem działania.

Ponieważ przesłanka (1) w postaci uogólnionej, jaką zaproponowałem wyżej, nie uwzględnia wszystkich logicznie możliwych zbiorów w partycji Y , można więc co najwyżej twierdzić, że wniosek (3) wynika logicznie z przesłanek (1) i (2) *warunkowo* lub – używając dziś rozpowszechnionej ter-

²¹ W tym punkcie dostrzegam pewną zbieżność z koncepcją wyjaśniania przyczynowego M. Webera z artykułu „Die Objektivität sozialwissenschaftlicher und sozialpolitischer Erkenntnis”: „[...] typ idealny pełni jedną tylko funkcję w badaniach empirycznych. Tą funkcją jest porównywanie go z empiryczną rzeczywistością, aby ustalić różnice i podobieństwa, aby opisać ją przy pomocy jak najbardziej jednoznacznych pojęć, zrozumieć ją i wyjaśnić przyczynowo”. (Weber 1904, cyt. za: Kmita i Nowak 1968, 67)

²² Realizacja w tym miejscu postulatu Feyerabenda, aby w wyjaśnianiu odgraniczyć „znaczenie zdania obserwacyjnego” od „przyczyny wytworzenia” go, wymagałaby przyjęcia nieredukcyjnej analizy przyczynowości; por. (Feyerabend 1962/1979, 136). To zadanie wykracza jednak poza przyjęte tu ramy rozważań.

minologii – *ceteris paribus* względem wiedzy, która jest wykorzystana w wyartykułowaniu zwłaszcza przesłanki (2).

Przywołajmy dyskutowany w Punkcie 4.2 hipotetyczny przykład Rubina. Uwzględnienie tylko fragmentu danej populacji prowadzi do wniosku, że nowa metoda operacyjna jest lepsza, gdyż wszyscy jej poddani żyją dłużej, leczeni starą metodą. Jednak uwzględnienie dodatkowych informacji – bądź to obserwacji w pozostałej części populacji, bądź to kontrfaktycznych (jaka byłaby długość ich życia, gdyby zostali poddani drugiej z metod operacyjnych) w odniesieniu do tej samej grupy – prowadzi do przeciwnego wniosku. Ten przykład ilustruje więc, że uwzględnienie dodatkowej informacji w przesłankach wnioskowania wyjaśniającego może prowadzić do zmiany wniosku.²³

Wyjaśniane zjawisko jest w przypadku ogólnym – a więc niezależnie od przypadków, gdy możliwe jest wnioskowanie *tollendo ponens* – zaliczane do tego z elementów partycji (1), który – wzięwszy pod uwagę informacje, jakie uzyskujemy w przesłance (2) – ma największą wiarygodność. Wśród wnioskowań probabilistycznych Neyman właśnie maksymalizację wiarygodności interpretował jako wnioskowanie najbliższe *tollendo ponens*.²⁴ W szczególności, gdy wiarygodność jednej z alternatyw zbliża się do jedności. Zdaniem Neymana, świadczy to bowiem o tym, że

[...] nie istnieją inne alternatywne hipotezy, których przyjęcie robiłoby zaobserwowany fakt w znaczniejszym stopniu bardziej prawdopodobnym, niż jest on po przyjęciu hipotezy sprawdzanej. (1930, 9-10)

Zaproponowana tu koncepcja subsumpcji proceduralnej, nie niwecząc odrębności koncepcji wyjaśniania subsumpcyjnego i przyczynowego, pozwala – jak sądzę – na zadowalającą charakterystykę związku subsumpcji. Przytaczając w wyjaśnieniu uogólnienie proceduralne, oparte na modelu przyczynowym, wyodrębniamy pewną dziedzinę rzeczywistości, do której ono się odnosi. Subsumpcja, czyli podporządkowanie wyjaśnianego zdarzenia prawidłowości wyrażonej w uogólnieniu projekcyjnym, w ujęciu proceduralnym polega więc zasadniczo na pokazaniu związku tego zdarzenia z dziedziną obowiązywalności uogólnienia projekcyjnego. Pokazanie, że wyjaśniane zdarzenie należy do dziedziny, w której obowiązuje uogólnienie projekcyjne, jest wystarczające do pokazania, że jest ono elementem prawidłowości wyrażonej w tym uogólnieniu.

²³ Szczegółową dyskusję można także znaleźć w (van Fraassen 1980, 114-18).

²⁴ Por. też (Backe 1999).

W niektórych przypadkach taką właśnie rolę odgrywa związek wynikania logicznego. Wówczas można uznać go za przypadek zależności subsumpcji proceduralnej. Nie jest to jednak jej warunek konieczny. Spóbję wyjaśnić to systematycznie. Hempel, wprowadzając wymóg subsumpcji jako związku wynikania logicznego, brał pod uwagę szczególną postać praw przyrody. Schemat praw przyrody był następujący: miały to być zdania wyrażające zachodzący empirycznie związek pomiędzy elementami dwóch zbiorów przedmiotów, a raczej zbioru nadrzędnego i podrzędnego. Związek podporządkowania między jednym zbiorem a drugim nie mógł wynikać logicznie z własności elementów podzbioru, gdyż byłaby to trywializacja pojęcia wyjaśniania. Istotne jest jednak to, że taka forma prawa przyrody zmuszałaby nie tylko do wyrażenia pewnej prawidłowości empirycznej, jaką obserwuje się pomiędzy elementami obu zbiorów, ale także do określenia własności, na podstawie których oba zbiory zostały wyodrębnione. Musiałyby być więc to własności wystarczające zarówno do należenia do danego zbioru przedmiotów, jak i zachodzenia prawidłowości wyrażanej w danym prawie. Opisanie zatem zdarzenia, które miało być wyjaśniane, w kategoriach tych własności nieuchronnie prowadziło do wniosku, że ma ono te własności, jakie przypisuje mu się w eksplanandum.

Być może to połączenie przez Hempela dwóch wymogów w schemacie prawa przyrody ma pewne korzenie w polemice logicznego pozytywizmu z metafizyką. W takim razie można jedynie zauważyć, że po oddzieleniu zadania identyfikacji zależności przyczynowych od zadania ich filozoficznej eksplikacji, nie ma filozoficznych podstaw do tego, by łączyć oba te wymogi, jak zrobił to Hempel w koncepcji DN i jej pochodnych.

Zwracając zresztą uwagę na kryterium proceduralne, założenie Markowa, zależność dla zmiennych kontrfaktycznych czy inne kryteria przyczynowe stosowane współcześnie w modelowaniu zależności przyczynowych, nie dostrzegamy połączenia tych dwóch wymogów w stylu Hempela. W żadnym bowiem z tych założeń nie wprowadza się wymogu określenia wystarczających warunków należenia do danej klasy przedmiotów, dla których zachodzi określona pewnymi wymogami statystycznymi – czy jak w przypadku kryterium proceduralnego także wymogami dotyczącymi interakcji z badaną rzeczywistością – prawidłowość.

Powróćmy na koniec tych rozważań do przykładu wyjaśnienia przyczyny cholery przez Snowa. W uzasadnianym przez niego modelu uogólnienie projekcyjne wyodrębnia binarną partycję z uwagi na wprowadzanie do organizmu wody zanieczyszczonej bakteriami cholery. W przypadkach, kiedy

taka binarna partycja nie pozwalała na podanie wyjaśnienia, Snow sugerował istnienie bogatszego modelu z uwagi na „inne czynniki”, które nie zostały przez niego uwzględnione. Znane były przypadki, gdy – świadome lub nieświadome – wprowadzenie bakterii cholery do organizmu, bądź nie wywoływało objawów chorobowych, bądź tylko słabe (Snow 1936, 111-13).

Snow w każdym przypadku podawanego przez siebie wyjaśnienia starannie wykluczał inną możliwość zachorowania na cholere niż przez wprowadzenie do układu pokarmowego bakterii cholery. Badał więc wszystkie okoliczności – które za istotne uznaliby zwolennicy koncepcji alternatywnej do miazmatycznej – aby pokazać, że z uwagi na nie zachodziła żadna różnica między tymi, którzy padli ofiarą epidemii, a tymi, którzy jej uniknęli. Tym samym więc przeprowadzał rozumowanie wykluczające przynależność danego przypadku do innej z możliwych i relewantnych alternatyw niż ta, że chorobę w tym przypadku spowodowało wprowadzenie do organizmu bakterii cholery wraz z wodą lub pokarmem.

Starannie w każdym z wyjaśnianych przypadków Snow starał się określić, że faktycznie zaistniał kontakt z powodującym cholere mikroorganizmem. W opisie wyjaśnianych przypadków zachorowań na cholere Snow w miarę swoich możliwości stosował pomiar, jakim było np. badanie składu chemicznego wody spożywanej przez osoby chore na cholere lub zmarłe z jej powodu. W przypadkach, gdy taki pomiar nie był możliwy, przywoływał możliwie pełny opis poszczególnych etapów procesu prowadzącego do wprowadzenia bakterii cholery do organizmu, jak np. kontakt z odzieżą chorych lub żywnością przetrzymywaną w ich pobliżu, przygotowywanie pokarmów przy niezachowywaniu podstawowych zasad higieny czy poszczególne etapy zanieczyszczania lokalnych źródeł wody pitnej.

Snow starał się wyjaśniać także korelację zachorowalności z określonymi zawodami (zwł. górniczy), rodzajem gleby, statusem społecznym i związanym z nim niskim poziomem higieny oraz złymi warunkami mieszkaniowymi) oraz z dzielnicą miasta. W zasadzie jednak, poza ostatnim przypadkiem, te wyjaśnienia nie spełniają wprowadzonych tu wymogów modelu proceduralnego. Przede wszystkim z uwagi na to, że Snow nie dysponował pełnym modelem przyczynowym, który pozwalałby na wygenerowanie odpowiednich uogólnień projekcyjnych.²⁵

²⁵ Innym przykładem błędnego wyjaśnienia było tłumaczenie przez Snowa sezonowości nasilania się epidemii cholery obyczajami kulinarnymi, np. Anglików czy Szkotów. Nie brał on pod uwagę możliwości – bliższej koncepcji miazmatycznej – że ta sezonowość związana jest ze zmianami temperatury wody w Tamizie oraz towarzyszącymi temu zmianami w popu-

Zakres stosowalności modelu przyczynowego Snowa jest więc wąski i nie pozwala na generowanie uogólnień projekcyjnych, które miałyby odniesienie do szerszego kontekstu niż system dystrybucji wody w południowych dzielnicach Londynu. Poważne ograniczenia tego modelu, związane np. z nieuwzględnieniem różnic osobniczych (np. poziomu odporności czy pewnych własności układu pokarmowego), zależności nasilenia cholery od klimatu czy istnienia różnych odmian bakterii cholery, nie zmieniają jednak podstawowego faktu, że w zakresie swojej stosowalności spełnia on wymogi proceduralnej koncepcji wyjaśniania.

9.4 Projekcyjność w proceduralnej koncepcji wyjaśniania

Na zakończenie warto podjąć problem, na jaki zwrócono uwagę podczas dyskusji podejścia BN do modelowania zależności przyczynowych. Z punktu widzenia prowadzonych badań empirycznych, istnieje niekiedy trudność z jednoznacznym wskazaniem modelu przyczynowego. Naukowiec niekiedy nie dysponuje wystarczającymi danymi do wykazania, że spełnione jest kryterium proceduralne dla postulowanego modelu. Często konstruowanie modelu rozpoczyna się od dokonania oszacowania rozkładu dla zmiennych obserwowalnych danego modelu. Przy braku danych manipulacyjnych nie można jednoznacznie stwierdzić, czy znany jest pełny zbiór zmiennych modelu. Istnieje jednak wiele modeli, które są zgodne z rozkładem określonym wyłącznie dla zmiennych obserwowalnych tego modelu. Różnice między tymi modelami sprowadzają się do odmiennego określania zależności przyczynowych oraz występowania innych zmiennych ukrytych (nieobserwowalnych). Jak podkreśla Pearl: „bez żadnych ograniczeń na rodzaj modeli branych pod uwagę – naukowiec nie będzie w stanie określić struktury badanych zjawisk” (2000, 45).

W wielu omówionych w poprzednim rozdziale koncepcjach wyjaśniania *implicite* przyjmowano założenie, że można *jednoznacznie* zidentyfikować poprawne wyjaśnienie. Trudności, jakie dyskutowałem w odniesieniu do koncepcji Salmona – część z nich uznał on sam – wyraźnie jednak wskazują – jak skonkludował to Woodward – że jest to zbyt wygórowane oczekiwanie

lacji mikroorganizmów, które mogłyby być nośnikami bakterii cholery w wodzie. Model przyczynowy Snowa nie pozwalał na uwzględnienie takiej zależności. Ponadto Snow nie brał pod uwagę możliwości, że bakteria cholery może być uzależniona od cyklu biologicznego organizmów żyjących w wodzie.

i należałoby je raczej zastąpić odniesieniem do pewnej klasy alternatywnych koncepcji.²⁶

W swojej krytyce algorytmów, które Spirtes, Glymour i Scheines zaproponowali, jako mechanizm generujący bayesowskie modele przyczynowe z danych nieeksperymentalnych, Freedman podkreśla, że: „te algorytmy prowadzą do estymacji nie modeli liniowych, lecz klas równoważnych modeli liniowych” (1997b, 180-181). Różnica jest zasadnicza, gdyż o tych drugich nie można jednoznacznie twierdzić – na co zwróciłem wcześniej uwagę za Pearlem – że w każdym przypadku poprawnie została zidentyfikowana zależność przyczynowa (Freedman 1997a, 144-145).

Spśród różnych analiz Freedmana, wykazujących niejednoznaczność wniosków przyczynowych na podstawie danych doświadczenia, najbardziej przekonujące – jak sądzę – jest studium związku przyczynowego między spożyciem soli a ciśnieniem krwi. Freedman wraz z D. Petitti analizują „hipotezę dotyczącą soli”, którą zwięźle wyrażają: „Powszechnie sądzi się, że spożycie soli prowadzi do podwyższonego ciśnienia krwi oraz większego ryzyka ataku serca czy udaru” (2002, 1). Jej konsekwencją jest, że spożycie soli powinno być drastycznie ograniczone. Szczegółowo analizując podstawowe rodzaje dowodów, które mają potwierdzać „hipotezę dotyczącą soli”, a mianowicie eksperymenty na zwierzętach i ludziach oraz badania obserwacyjne na ludziach (tzw. program Intersalt przeprowadzony w 52 krajach), wykazują, że dowody te nie dają wystarczającego potwierdzenia tej hipotezie. I tak na przykład, w przypadku osób młodych (tj. do 20 roku życia) korelacja wydaje się odwrotna (choć nie w 4 egzotycznych spośród 52 krajów), a efekt dla osób o normalnym ciśnieniu krwi niezauważalny. Jeśli zaś chodzi o wpływ na zdrowie, to daje się on zauważyć zwłaszcza u osób otyłych, ale i tak jest różnokierunkowy.

Ogólny wniosek zatem brzmiałby, że istnieje wiele modeli, które są obserwacyjnie zgodne z danymi, jakich dostarczają dotychczasowe badania tej hipotezy. Wśród nich znalazłby się także i model zgodny z „hipotezą dotyczącą soli”, ale tylko po odpowiednim przekształceniu, tj. po zawężeniu wyjaśnianego obszaru zjawisk.

²⁶ S. Kamińskiego studium wyjaśniania, zwłaszcza przyczynowego, potwierdza takie rozwiązanie: „Trudno [...] byłoby to zrobić [tj. wskazać jednoznacznie właściwą przyczynę] w wyjaśnianiu rzeczywistości wziętej w aspekcie kwalitatywnym lub kwantytatywnym, chyba że w wyjątkowo prostej sytuacji. Zawsze bowiem zachodzi ogromne bogactwo (wielość i różnorodność) czynników sprawczych i warunków”. (1966/1989, 175)

Wobec wyników analiz Freedmana oraz wcześniej dyskutowanych rozwiązań zaproponowanych przez Rubina w odniesieniu do przykładu skuteczności dwóch metod operacyjnych nasuwa się następujące rozwiązanie problemu występowania równoważnych obserwacyjnie modeli zależności przyczynowych. Bez przyjęcia dodatkowych założeń dostępne jest jedno z dwóch rozwiązań. Można – jak proponują to Glymour i współpracownicy – ograniczyć się do określenia klasy modeli, które są równoważne obserwacyjnie, a następnie postawić pytanie, czy wszystkie modele w tej klasie udzielają jednoznacznej odpowiedzi na pytanie o zachodzenie zależności przyczynowej między dwoma zjawiskami.

Inne rozwiązanie, które w swoich analizach przyjął Freedman, to wyodrębnienie podpopulacji z uwagi na pewien dodatkowy parametr, który nie został uwzględniony w modelu. Dla każdej z tych populacji otrzymamy następnie jednoznaczny model zależności przyczynowych.

Można wreszcie, jak dopuszczają to metody modelowania BN oraz RCM, przyjąć dodatkowe założenia. W omówionym szczegółowo w Rozdziale 5. przykładzie z oceną skuteczności starej i nowej metody operacyjnej założeniem, jakie przyjął Rubin, było określenie rozkładu dla populacji. Względem niego właśnie obliczane następnie były efekty kontrfaktyczne zastosowania drugiej operacji w stosunku do pacjentów z danej grupy. W ten sposób otrzymuje się jeden wspólny model dla obu grup.

Zarówno klasycy teorii wyjaśniania, jak Hempel, jak i nowsi autorzy, zwłaszcza Woodward, zgodnie wydają się ukrycie przyjmować tezę o naturalistycznych implikacjach teorii wyjaśniania. Ponieważ wyjaśniamy w oparciu o obserwowane regularności lub efekty manipulacji, więc wyjaśnienie nie może prowadzić do stwierdzenia istnienia obiektów innych niż znane z doświadczenia lub z jego naukowych teorii. Hempla wymóg empirycznej testowalności, Salmona żądanie śledzenia czasoprzestrzennej ciągłości procesów fizycznych czy Woodwarda żądanie manipulowalności implikują pewne ograniczenia ontologiczne w stosunku do tego, co może występować jako eksplanans oraz eksplanandum.

W dynamice wyjaśniania, którą nakreśliłem dla modelu proceduralnego, nie ma podstaw do wyprowadzania takiego rodzaju implikacji naturalistycznych.²⁷ Jeżeli tylko potrafimy określić model zależności przyczynowych,

²⁷ Gdyby rozważać tezę, że z uwagi na rolę założenia Markowa w modelach BN proponowana niżej koncepcja wyjaśnienia jest szczególnego rodzaju wnioskowaniem zmierzającym do najlepszego wyjaśnienia (Grobler 1991, 177-80), to należałoby uwzględnić taką kwalifika-

wówczas także można sformułować na jego podstawie uogólnienie projekcyjne i wyjaśnienie. W tym miejscu nie podejmuję dyskusji z tezą naturalizmu, sygnalizuję jedynie, że zaproponowany tu model – w odróżnieniu od większości proponowanych współcześnie – nie zakłada milcząco słuszności tezy naturalistycznej.²⁸

Analogicznie można argumentować, że istnienie modeli przyczynowych równoważnych obserwacyjnie nie implikuje „metafizycznego pluralizmu nomologicznego”, który Cartwright rozumie jako „doktrynę, że naturą rządzą różne systemy praw w różnych dziedzinach, niekoniecznie ze sobą powiązane w jakiś systematyczny czy jednorodny sposób; składanka praw (ang. patchwork of laws)” (1999, 31).

W Punkcie 9.1 szczegółowo pokazałem, że model wyjaśniania zaproponowany przez Woodwarda jako warunek konieczny wprowadza uwzględnienie (kontrfaktycznych) informacji o efektach manipulacji danym zjawiskiem, jego warunkami początkowymi lub brzegowymi. Każdy model wyjaśniania, który respektuje odmienną koncepcję wyjaśniania przyczynowego, musi dopuszczać taką możliwość, jaką wyeksponował w swoim modelu Woodward. Nie sądzę jednak, aby podanie informacji związanych z efektami manipulacji było warunkiem koniecznym wyjaśniania. Przyjęcie takiego postulatu prowadzi Woodwarda do rozciągnięcia pojęcia interwencji z jednostkowych interakcji na interakcje czysto hipotetyczne, wskutek czego trudno jest je odróżnić od związku czysto myślnego między eksplanansem a eksplanandum.

Można, jak Psillos (2002, 105-106), krytykować ten element modelu Woodwarda, pokazując, że niezmienniczość na interwencje jest tylko przejawem występujących w wyjaśnieniach uogólnień, a nie ich istotną charakterystyką. W związku z tym żądanie, by informacje o manipulacjach były elementem koniecznym wyjaśnień, nieuchronnie prowadzi do kołowości w konstrukcji koncepcji manipulacyjnego modelu wyjaśniania.

Inny sposób podważenia tego elementu w koncepcji Woodwarda to przywołanie przykładów historycznych, które są z nią niezgodne. Sądzę, że znakomitą ilustracją jest różnica stanowisk między komisją the Board of Health,

cję, istotnie ograniczającą zastosowania tego rodzaju wnioskowań. Z uwagi na obszerność tej problematyki, pozostawiam ją do osobnego studium.

²⁸ Inna różnica, której tu nie dyskutuję szczegółowo, to dopuszczalność wyjaśnień przyczynowych dla zdarzeń jednostkowych w bayesowskich modelach przyczynowości (Pearl 2000) zgodnie z postulatem D. Davidsona (1967/2001, 161-162). Por. też (Mellor 1995, 130-131).

działającą z ramienia parlamentu brytyjskiego, a stanowiskiem Snowa w sprawie przyczyn cholery oraz sposobów jej zapobiegania. Zaproponowane przez Snowa wyjaśnienie tych przyczyn, mimo ograniczeń, a niekiedy i błędów, można uznać za zadowalające wyjaśnienie w świetle modelu proceduralnego. Koncepcja miazmatyczna komisji, według której przyczyny choroby są w losowo generowanej nieorganicznej truciznie cholery, nie spełnia tych wymogów przede wszystkim z uwagi na brak danych manipulacyjnych i niespełnianie kryterium proceduralnego.

Jak podkreślają zgodnie współcześni komentatorzy tej dyskusji między komisją a Snowem (Paneth i in. 1998), swoją trwałość – mimo braku adekwatnego wyjaśnienia jej przyczyn – miazmatyczna koncepcja cholery zawdzięczała skuteczności w jej zapobieganiu. Nawet podane przez Snowa zalecenia co do zapobiegania i zwalczania cholery pozostawały pod wpływem praktycznych zaleceń zwolenników koncepcji miazmatycznej.

Ten fragment dziejów badań epidemiologicznych dostarcza więc jednocześnie dwóch przykładów niezgodnych z wymogami koncepcji manipulacyjnej Woodwarda. Snow podał zadowalające wyjaśnienie przyczyn cholery, mimo braku jasno określonego sposobu wpływania na tę chorobę. Z drugiej strony natomiast komisja, posiadająca błędną koncepcję przyczyn tej choroby, trafniej wskazywała sposoby jej zapobiegania.

Określenie (kontrfaktycznych) informacji o efektach manipulacji eksplanandum wedle powyższych przykładów nie wydaje się zatem ani elementem koniecznym, ani wystarczającym podania zadowalającego wyjaśnienia. To, że takie informacje mogą wystąpić w podanym wyjaśnieniu, jest zgodne z modelem proceduralnym, jednak jest traktowane jako przejaw, a nie element konstytuujący zadowalające wyjaśnienie.

Na zakończenie chciałbym podkreślić pewną własność modelu proceduralnego wyjaśnienia, która może stać się źródłem nieporozumień. Niezależnie od tego, czy wyjaśnienia traktuje się jak argumenty (Salmon 1998, 102-104), intuicyjne rozumienie tego pojęcia wiąże się z pewnego rodzaju wnioskowaniem. To z kolei może prowadzić do pozornej trudności, że podanie modelu nie jest związane z wnioskowaniem, a więc nie może być adekwatnym wyjaśnieniem. Przypomnę więc to, co zostało wyraźnie przedstawione w Rozdziale 4. – modele przyczynowości są *jednocześnie* mechanizmem inferencyjnym, który pozwala przeprowadzać wnioskowania przyczynowe, a w przypadku modeli identyfikowalnych także sprawdzać zgodność tych wniosków z danymi doświadczenia.

Podstawy takiego nieporozumienia w przypadku ogólnym usuwa R. Wójcicki:

Istnieje wyraźne pokrewieństwo między pojęciem modelu oraz pojęciem prawa. Każde prawo jest w istocie rzeczą modelem zjawiska, którego dotyczy z uwagi na ten typ prawidłowości, pod który podpada prawidłowość opisywana prawem. (1974, 300)

Zakończenie

Metodologia proceduralna – między realizmem a antyrealizmem naukowym

Współczesne filozoficzne koncepcje nauki pozostają w cieniu prac logicznych empirystów, zwłaszcza Rudolfa Carnapa i Hansa Reichenbacha. Są ich kontynuacją bądź zanegowaniem. Kontynuacją unikającą wielu podstawowych trudności wcześniejszych wersji empiryzmu jest empiryzm konstruktywny Basa van Fraassena. Stanowisko to systematycznie rozwija dwie tezy:

Nauka zmierza do podania nam teorii, które są adekwatne empirycznie, a akceptacja teorii wiąże się z przekonaniem tylko o jej adekwatności empirycznej. (van Fraassen 1980, 12)

Teoria jest adekwatna empirycznie, gdy „zachowuje zjawiska”, a więc gdy ma przynajmniej jeden model, który zawiera *wszystkie* zjawiska, także i te dotąd niezobserwowane. Akceptacja teorii jest natomiast czymś więcej niż przekonanie o jej prawdziwości lub adekwatności empirycznej. Ta różnica ujawnia się przede wszystkim w przypadku, gdy naukowiec decyduje się na akceptację teorii z grona teorii równie adekwatnych empirycznie i jest gotowy na konfrontowanie zjawisk, a w przyszłości będzie opierał się na tej właśnie teorii.

W opozycji do empiryzmu, a w szczególności empiryzmu konstruktywnego, pozostają stanowiska realizmu naukowego, szeroko reprezentowane we współczesnej filozofii nauki i metodologii.¹ Za nadrzędny cel nauki uznaje się w nich podanie teorii naukowych, które są prawdziwe także w zakresie tego, co nieobserwowalne,² czyli tego, co wykracza poza kryterium adekwatności empirycznej.

Putnam przeniósł dyskusję nad realizmem naukowym na grunt pragmatyzmu:

Negatywne argumenty sprowadzają się do pokazania, że różne redukcyjne czy operacjonalistyczne filozofie są po prostu nieudane. [...] Pozytywny argument za realizmem tkwi w tym, że jest to jedyne stanowisko, które nie sprowadza sukcesu nauki do cudu. W tym, że terminy dojrzałych teorii naukowych mają zwykle odniesienie przedmiotowe (jak sformułował to Richard Boyd), w tym, że teorie zaakceptowane w dojrzałej nauce są zwykle przybliżenie prawdziwe, w tym, że ten sam termin może odnosić się do tej samej rzeczy, nawet gdy występuje w różnych teoriach – te twierdzenia realista naukowy uważa nie za prawdy konieczne, lecz za część jedynego naukowego wyjaśnienia sukcesu nauki, a zatem i za część każdego adekwatnego opisu nauki i jej relacji do badanych przez nią przedmiotów. (Putnam 1975/1999, 319-20)

Nie wchodząc w subtelne różnice między odmianami realizmu naukowego³, przypomnę trafne stwierdzenie Judyckiego, że argumenty odwołujące się do sukcesu nauki⁴ są kategorialnie rozbieżne z celem, jakim dla realisty jest wykazanie realności nieobserwowalnego świata teorii naukowych:

Nie pomoże tu też pragmatyzm w postaci odwołania się do faktu sukcesu eksplanacyjnego, prognostycznego i technologicznego naszej nauki. W świecie zjawisk możemy dużo ‘wiedzieć’ i działać sensownie, lecz w problemie poznania ludzkiego chodzi o to, czy świat, z którym mamy do czynienia i w którym działamy, jest tylko zjawiskiem, czy też ‘światem samym w sobie’. (Judycki 2005, 147)

¹ Feyerabend jako zasadniczą alternatywę wprowadza tu opozycję między realizmem (naukowym) a instrumentalizmem (1964/1979). Taka alternatywa może być szczególnym przypadkiem przyjętego tu odróżnienia. Por. też (Wójcicki 1991, 1-3).

² Szczegółową analizę kluczowego tu pojęcia obserwowalności przeprowadził Przełęcki (1988, 43-44).

³ Por. (Kałuszyńska 1995, 15-20; Niiniluoto 1999; Psillos 1999).

⁴ Dotyczy to także twierdzeń o zbieżności na gruncie bayesianizmu; por. (Dorling 1992; Juhl i Kelly 1994).

Mimo że pragmatyczne przywołanie sukcesu nauki nie stanowi wystarczającej podstawy do wykazania realności nieobserwowalnych przedmiotów świata nauki, zgadzam się z Niiniluoto, że wymaga on pełniejszego wyjaśnienia niż to, które mogą zaproponować empiryści konstruktywni,⁵

którzy zachowują klasyczne pojęcie prawdy nawet dla teorii naukowych, ale uznają je za całkowicie nieistotne w analizie postępu naukowego. Sugerują oni, że nauka *jest* skuteczna w praktyce (w formułowaniu prawdziwych obserwacyjnych predykcji lub w rozwiązywaniu problemów empirycznych), ale ten fakt nie wymaga żadnego wyjaśnienia. Van Fraassen zauważa, że nic dziwnego w tym, że teorie ‘działają’, gdyż *wybieramy* te teorie, które ‘prze-trwają’ ‘zacieklą rywalizację’. [...]

Te wymówki uznaję za niewystarczające. Rozważmy analogię: dlaczego nasze samochody są szybsze niż pięćdziesiąt lat temu? Nie wystarczy powiedzieć, że kupujemy dziś szybsze samochody niż dawniej (jak twierdziłby van Fraassen) [...]. Nasze pytanie o wyjaśnienie wymaga podania stosunkowo trwałej możliwości samochodu, dzięki której jest szybszy. Musimy wskazać na pewną własność (jak struktura jego silnika), która powiąże zachowanie tego artefaktu z jego funkcjonowaniem w jego naturalnym środowisku. (Niiniluoto 1999, 197-98)

Taką brakującą „własność”, jak staram się to poniżej uzasadnić, podaje metodologia proceduralna. Przedtem jednak proponuję naświetlenie różnicy stanowisk realizm-antyrealizm w kontekście przykładu badań Snowa nad przyczynami cholery.

Przyjmę, że najprawdopodobniej Snow był przekonany o istnieniu procesu czy mechanizmu przyczynowego, który powoduje, że jeśli ktoś pije skażoną wodę, to z dużym prawdopodobieństwem stanie się jedną z ofiar cholery. Można postawić dwa⁶ pytania: (1) czy takie przekonanie jest prawdziwe oraz (2) czy stwierdzona przez Snowa na podstawie obserwacji zależność między zmiennymi (skażona woda oraz zachorowalność na cholere) jest zależnością przyczynową.

Prymitywny realista naukowy powiąże oba pytania, traktując odpowiedź na pytanie (1) jako warunek udzielenia odpowiedzi na pytanie (2). Empirysta prymitywny na oba pytania odpowie negatywnie. Jak staram się to pokazać poniżej, proponowana tu metodologia proceduralna pozwala zintegrować słuszne tezy obu tych stanowisk i udzielić odpowiedzi pozytywnej na pytanie (2), odmawiając nauce kompetencji do odpowiedzi na pytanie (1).

⁵ Por. też (Kałuszyńska 2001, 78-79).

⁶ Odpowiednio do wprowadzonego we Wstępie rozdziału kompetencji w badaniu przyczynowości na kwestie filozoficzne i kryterialne.

Spośród bogatej palety stanowisk realistycznych w ogólnej metodologii i filozofii nauki zwróćmy uwagę na stanowisko socjologa z Nuffield College w Oxfordzie H. Goldthorpe'a (2000; 2001). Wyartykułował on na gruncie metodologii badań społecznych stanowisko realizmu naukowego w kwestii ustalania przyczyn. Zasadniczy dla tego stanowiska jest postulat określenia mechanizmu przyczynowego, który – choć nieobserwowalny – generuje obserwowane zjawiska, na podstawie których dokonuje się stwierdzenia przyczynowego powiązania tych zjawisk.⁷

Podkreślając, że odejście od empiryzmu w metodologii – i związane z tym przejście w kierunku metod jakościowych – nie rozwiązało problemu metod ilościowych, a wzbogaciło je o nowe, autor *On sociology* projektuje taką metodologię badań społecznych, która zintegrowałaby zalety podejścia jakościowego i ilościowego. Zasadniczy jej rys tkwi w połączeniu ilościowych, statystycznych metod analizy danych z jakościową narracją teoretyczną, która wyjaśni zaobserwowane zależności, postulując generujący je proces przyczynowy.⁸ Ze względu na teoretyczną preferencję Goldthorpe'a, skłaniającą go w kierunku teorii racjonalnego działania, podającej wyjaśnienia na poziomie mikro, a więc na poziomie indywidualnych działań członków danej społeczności, warstwę teoretyczną proponowanej metody badań wiąże on z teorią racjonalnego działania. Postulowane procesy przyczynowe nie są „bezpośrednio obserwowalne” oraz „podawane wyjaśnienia tych procesów przyczynowych, mające na celu ujaśnienie (ang. illuminate) «czarnej skrzynki», obecne w czysto statystycznej analizie, nigdy nie są ostateczne”, gdyż mimo potwierdzenia w testach empirycznych, nie można wykluczyć możliwości bardziej subtelnej wyjaśnienia (2000, 149). Jako wzorcowy przykład takiego właśnie postępowania Goldthorpe podaje badania przeprowadzone nad zależnością między paleniem papierosów a rakiem płuc. Konkluzja jest następująca:

Analiza statystyczna danych obserwacyjnych wykazała silną asocjacje między paleniem a rakiem płuc oraz że jest ona odporna na wprowadzenie szeregu czynników, które mogłyby okazać się „wspólną” przyczyną obu tych zjawisk. Jednak kluczowe dla twierdzenia, że jest to związek przyczynowy, było opracowanie leżącego u podstaw generującego tę zależność procesu na bazie wy-

⁷ W mniej wyartykułowanej postaci tę koncepcję proponują także Cox (1992, 297), Cox i Wermuth (1996, 220-21) oraz Simon i Iwasaki (1988).

⁸ Zgodnie z koncepcją Goldthorpe'a, badanie społeczne przebiega w trzech etapach: ustalenie zjawisk, które stanowią eksplanandum, wprowadzenie hipotezy o generującym procesie przyczynowym na poziomie działania społecznego, a następnie testowanie tej hipotezy (2000, 151-60).

izolowania znanych czynników rakotwórczych w dymie papierosowym, dowodów histopatologicznych z nabłonka oskrzeli palaczy itp. (Goldthorpe 2000, 150).

Nie jest jednak do końca jasne, czy zasadniczą kwestią jest *rzeczywiste* istnienie tego procesu przyczynowego, czy raczej *teoretyczna* hipoteza o istnieniu takiego procesu, jako konieczne uzupełnienie statystycznych analiz zależności między zmiennymi. W dalszym komentarzu Goldthorpe podkreśla, że

przyczyna nie może być ustalona po prostu przez ogólne procedury wnioskowania statystycznego bez dodatkowej wiedzy o przedmiocie badania. Jeśli wymagane jest wyjaśnienie procesów, o których sądzi się, że tworzą statystycznie udowodnioną zależność, wówczas takie wyjaśnienie musi być podane zasadniczo na podstawie wiedzy o przedmiocie badania, a im dogłębniej takie wyjaśnienie jest przesycone dominującą teorią, tym bardziej jest spójne i testowalne. (Goldthorpe 2000, 150)

Zmiana teorii musi zatem implikować także zmianę postulatu o procesie generującym zaobserwowane zależności. Zrównoważenie przeważa więc zdecydowanie w stronę teoretycznego charakteru postulatów o istnieniu procesów przyczynowych leżących u podstaw danych. W sposób charakterystyczny dla realizmu naukowego Goldthorpe implicite wyciąga z tego konkluzję o realności takich procesów postulowanych hipotetycznie. Ze względu na zmienność teorii, prawdziwość przypisuje się jedynie przekonaniu o istnieniu zaobserwowanych zależności przyczynowych. Natomiast akceptacja hipotezy o generującym je procesie może być motywowana jedynie względami pozaepistemicznymi. Tym samym więc proponowany przez Goldthorpe'a projekt metodologii badań społecznych – wbrew deklaracjom – jest zasadniczo zgodny z metodologią proceduralną. O istnieniu zależności przyczynowych rozstrzyga się w oparciu o zaobserwowane zjawiska, natomiast rozstrzygnięcie kwestii istnienia postulowanych procesów przyczynowych zostaje poza obrębem epistemicznych kompetencji filozofii i metodologii nauki.

Różnicę między testowaniem zależności przyczynowej a testowaniem procesu przyczynowego Goldthorpe omawia na przykładzie konsekwencji, jakie dla dzieci ma rozwód rodziców (2000, 156-58). Ustalona w badaniach korelacja między rozwodem a kończeniem przez dzieci edukacji na najwcześniejszym możliwym etapie, może nie być zależnością przyczynową, gdyż możliwe są konfundery tej zależności, np. osobowość czy konflikt między rodzicami, które zarówno powodują niestabilność związku małżeńskiego,

jak i złe wychowanie dzieci. Nie można przeprowadzić eksperymentu z randomizacją, by potwierdzić lub wykluczyć te konfundery. Goldthorpe powątpiewa też w siłę wnioskowania kontrfaktycznego RCM, gdyż to samo małżeństwo nie może jednocześnie się rozwiąć i nie rozwiąć (Holland 1986b). Można co najwyżej w analizie uwzględniać czynniki, pod względem których różnią się poszczególne przypadki, tak by przybliżyć się maksymalnie w badaniu obserwacyjnym do kontekstu randomizacji, np. przez dopasowanie (dzieci, których rodzice się rozwiedli, dobiera się z dziećmi, których rodzice się nie rozwiedli, tak aby pod każdym innym względem zachować podobieństwo między nimi). Nie jest możliwe wyczerpanie wszystkich czynników, a tym samym otrzymanie konkluzywnej odpowiedzi. Ponadto, im więcej takich czynników się uwzględni, tym bardziej sceptyczne otrzymuje się wyniki analiz.

Testowanie hipotez o procesie generującym miałoby, zdaniem Goldthorpe'a, polegać na możliwie szczegółowym opisie każdego z nich, aby wyrazić różnice. Można by je następnie sprawdzić, porównując dzieci, które straciły rodzica wskutek rozwodu i wskutek śmierci, porównując rodzzeństwo, które przeżywa rozwód rodziców w różnym wieku, oraz te dzieci, które mają ojczyzna lub macochę. Taka procedura jednak wydaje się zastosowaniem zakwestionowanego wcześniej dopasowania.

Wydaje się więc, że ostatecznie projekt Goldthorpe'a sprowadza się przede wszystkim do podkreślenia, że „wyjaśnienia przyczynowe nie mogą zostać sformułowane wyłącznie w oparciu o metodologię statystyczną: wymagana jest wiedza o przedmiocie badań w postaci wiedzy ogólnej, a przede wszystkim teorii” (2000, 159), oraz do postulatu, by wiedza teoretyczna była wyrażana w kategoriach działania na poziomie jednostkowym (mikro).

Pierwszy z tych postulatów jest jednym z ważnych kryteriów proceduralności (por. Rozdział 2.) i jako taki stanowi jeden z ważnych składników metodologii proceduralnej. Drugi – pozostaje z nią zgodny, chociaż wymagałby niezależnej argumentacji.

Przechodząc do drugiego ze stanowisk, przypomnę, że prymitywny empiryzm – jeśli w ogóle uzna postawione przeze mnie pytania (1) i (2) za sensowne – na oba udzielił odpowiedzi negatywnej. Zasadniczym powodem jest to, że „obserwacja zjawisk nie wskazuje jednoznacznie na rzekome związki przyczynowe, które stałyby za nimi” (van Fraassen 1980, 2). Takie stanowisko ujawnia echa przebrzmiałego dziś stanowiska radykalnie antymetafizycznego, które – jak zaznaczyłem we Wprowadzeniu – dominowało w drugiej połowie XIX wieku.

Znacznie bardziej subtelny empiryzm konstruktywny pozostawi pytanie (1) poza obszarem pytań podejmowanych w nauce, nie negując przy tym możliwości podania odpowiedzi, która posiada wartość logiczną, a w szczególności odpowiedzi twierdzącej. Przyjęta przez Snowa hipoteza o istnieniu mechanizmu przyczynowego może być prawdziwa, jak chciałby Goldthorpe – jej prawdziwość jednak wykracza poza kryterium adekwatności empirycznej i pozostaje bez istotnego związku z celem nauki. Jakiej odpowiedzi udzieli się na pytanie (1) jest rzeczą nieistotną, gdyż ważniejsze jest to, że dzięki akceptacji mechanizmu przyczynowego Snow określił w sposób adekwatny empirycznie zależność między obserwowalnymi zjawiskami.

W odniesieniu natomiast do pytania (2) van Fraassen rozważa możliwość: „dlaczego by nie dopuścić przyczynowości w modelach, dlaczego traktować je gorzej niż fotony?” (van Fraassen 1993, 437), które – mimo że też nieobserwowalne – przyjmuje się w modelach fizycznych. Powód odpowiedzi negatywnej jest mało zasadniczy:

postrzegam to w ten sposób, że modele, które proponują nam naukowcy, nie zawierają struktury, którą możemy opisać jako rzekomo reprezentującą przyczynowanie bądź jako odróżniającą przyczynowanie od innych zdarzeń, które nie są przyczynowaniem. (van Fraassen 1993, 437-38)

Jak podkreśla on, tego rodzaju odróżnienie „wprowadza się poza teorią”, a więc i poza naukę; nawet jeśli, jak Cartwright (1993, 426), przywołać sposób wypowiedzania się naukowców. Van Fraassen podkreśla, że byłoby to jednak reifikowaniem terminów, które naukowcy czerpią z języka potocznego, poruszając się nie na gruncie nauki, lecz na gruncie jej aplikacji.

Dodaje on, iż tego rodzaju reifikacji towarzyszy „metafizyka Whiteheadowska” (Whitehead 1959), która wypowiedzi odwołujące się do zależności przyczynowych wiąże z istnieniem hierarchii poziomów. To, że ból głowy ustępuje po wzięciu przez mnie aspiryny, należy rozumieć tak, że moje działanie (wzięcie aspiryny) powodujące skutek (ustąpienie bólu głowy) jest „na powierzchni” tego, co dzieje się faktycznie. Moje działanie ma skutek dlatego, że aspiryna usuwa ból głowy (poziom niżej), a to jej działanie wynika z tego, co dzieje się na poziomie niżej. Obecność kwasu acetylosalicylowego powoduje, że naczynia krwionośne się rozszerzają. Dzieje się tak z kolei przez to, co dzieje się na jeszcze niższym poziomie reakcji chemicznych. Te z kolei można wyjaśnić – przechodząc na poziom niżej – przez zależności przyczynowe sformułowane dla cząstek, atomów, elektronów, protonów i kwarków. W konsekwencji:

Zdarzenia na powierzchni i ogólnie na wszystkich wyższych poziomach są przyczynowaniem w sensie wtórnym, *dzięki* przyczynowaniu zachodzącemu na poziomie najbardziej fundamentalnym. Najważniejszym pojęciem przyczynowości jest pojęcie zależności przyczynowej zachodzącej pomiędzy fundamentalnymi zdarzeniami fizycznymi. (van Fraassen 1993, 442)

Stanowisko empiryzmu konstruktywnego zdecydowanie oponuje przeciw narzucaniu nauce takiej hierarchii poziomów. Van Fraassen skłonny jest natomiast przyjąć pewną wersję poglądu R. Collingwooda. Zgadzając się ze sceptycyzmem Russella co do istnienia przyczynowości na najbardziej fundamentalnym poziomie fizyki (1912-1913), Collingwood hierarchię Whiteheadowską odwraca do góry nogami. W sensie pierwotnym więc języka przyczynowego używamy, mówiąc, że głowa przestała mnie boleć, bo wziąłem aspirynę. Ogólnie – język przyczynowy zachowuje swoje znaczenie tylko wtedy, gdy podmiot oznacza osobę. W pozostałych przypadkach jest używany metaforycznie.

Van Fraassen nie traktuje poglądów Collingwooda jako poszerzonej wersji stanowiska empiryzmu konstruktywnego: „Zdaję sobie sprawę, że idea Collingwooda nie jest niczym więcej niż tylko ideą. Nie zamierzam jej obecnie rozwijać” (1993, 433). Stanowi ona jednak taką alternatywę dla metafizyki Hume’owskiej i Whiteheadowskiej, jaka mogłaby posłużyć do odpowiedniej modyfikacji stanowiska empiryzmu konstruktywnego (1993, 439).

Powyżej przeprowadzona dyskusja stanowiska realizmu i empiryzmu w kwestii przyczynowości prowadzi do wniosku, że oba stanowiska w ostatecznym rozrachunku zbliżają się do stanowiska proceduralnego rozwijanego w tej książce. Projekt metodologii proceduralnej mieści się, jak sądzę, w dotąd na ogół omijanym przejściu między Scyllą empiryzmu a Harybdą realizmu. Jak w realizmie, traktuje się przyczynowość jako zasadniczą kategorię w badaniach nieeksperymentalnych, a jednocześnie, jak w empiryzmie, nie uzależnia się jej wprowadzenia od uznania, że ostateczną płaszczyzną ontologiczną jest świat fizyki fundamentalnej. Wynika to przede wszystkim z przypisania przyczynowości pierwotnego charakteru w metodologii badań nieeksperymentalnych.

Ze względu na ograniczenia tych stanowisk wprowadziłem pojęcie *interakcji badawczej* ze światem danym w doświadczeniu jako pojęcie pierwotne i nieredukowalne. Odwołanie się wyłącznie do zarejestrowanych stanów zjawisk (np. w postaci baz danych rejestrujących zmierzone wartości zmiennych) nie pozwala na wykazanie, że zależność między zmiennymi jest przyczynowa. Można przytaczać wiele filozoficznych argumentów, które

potwierdzałyby ten wniosek, ale zwróćmy uwagę przede wszystkim na rzecz następującą: jeśli dopuścić taki model zależności przyczynowych, w którym między każdymi dwoma z badanych zmiennych zachodzi zależność przyczynowa, to niezależnie od danych, jakimi dysponujemy, nie można wykazać, że zgodny z danymi jest tylko jeden model.⁹

Wskazanie na to, że zależność między zmiennymi jest przyczynowa, wymaga więc spełnienia *innego* jeszcze kryterium niż tylko zgodność z danymi. Takiego kryterium, jak pokazywałem wcześniej, nie dostarcza realista naukowy. Takim kryterium natomiast jest proponowana tu *interakcja* (przyczynowa) badacza z rzeczywistością. Wprowadzając to pojęcie jako pierwotne, unika się problemu kołowości w określaniu zależności przyczynowych między zmiennymi w kategoriach interakcji. W stosunku do proponowanego przez empirystę kryterium zgodności z doświadczeniem kryterium interakcji przyczynowej jest bardziej podstawowe. Określenie zgodności z doświadczeniem wymaga wcześniejszego zarejestrowania stanu badanych zjawisk, co jest możliwe tylko dzięki interakcji badacza z rzeczywistością.

Zwróćmy uwagę, że kryterium interakcji *pogłębia* intuicję, która stoi za kryterium zgodności z doświadczeniem, w kierunku interpretacji, jakiej żądałby realista naukowy. Kryterium interakcji, zgodnie z postulatem realizmu, wymaga istnienia *rzeczywistego* związku między badaczem a badaną rzeczywistością, ale zgodnie z postulatem empiryzmu ta interakcja dokonuje się i ujawnia się poprzez to, co jest dane w doświadczeniu.¹⁰

Jedną z konsekwencji tego pogłębionego sensu empirystycznego kryterium zgodności z doświadczeniem jest to, że – jak chcą realiści (Glymour 1985) – odróżnienie obserwowalne-nieobserwowalne ma przywrócone sobie właściwe miejsce.¹¹ Zgadzam się, że jest to podstawowe dla filozofii i metodologii nauk odróżnienie, aczkolwiek – wbrew empirystom – nie oczekiwałbym tu poważnych konsekwencji filozoficznych.

⁹ Szczegółowo ta kwestia jest dyskutowana w odniesieniu do bayesowskich modeli przyczynowych, gdzie każdy kompletny model dla zbioru zmiennych V (tworzony w sposób wyżej opisany) jest zgodny z danymi (Freedman 1997a).

¹⁰ Gdyby przyjąć sprawczą koncepcję przyczynowości na gruncie filozoficznym, wówczas można by nawet powiedzieć, że interakcja jest dana w doświadczeniu (rozumianym szerzej niż tylko to, co daje się zaobserwować); por. (Menzies i Price 1993).

¹¹ Jedną z konsekwencji nadania temu odróżnieniu szczególnego statusu epistemicznego w nauce jest zatarcie różnicy – w obrębie tego, co nieobserwowalne – między tym, co testowalne a nietestowalne empirycznie (Grobler 1991, 119).

Interakcja z badaną rzeczywistością D i wyróżnionymi zjawiskami V może bowiem prowadzić do efektów, które są obserwowalne w kategoriach zjawisk V' w D, ale nie w kategoriach zjawisk V. I tak Snow, ustalając wpływ skażonej wody na zachorowalność na cholere, podejmował interakcję z poszczególnymi przypadkami obecności w organizmie pacjenta nieobserwowalnej gołym okiem trucizny cholery. W tym sensie więc obserwowane przez niego zjawiska, jak woda skażona odchodami chorych, okres inkubacji bakterii w organizmie, patologia cholery czy zgon chorego, należą do innej kategorii zjawisk niż te, z którymi podejmował interakcję, badając mechanizm przenoszenia cholery i jej przebieg w organizmie chorego.

Oprócz opisanego mechanizmu przebiegu cholery w organizmie chorego (zakażenie następuje przez kanał oralno-fekalny), Snow wskazywał, jak nieobserwowalny dla niego mechanizm przenoszenia cholery łączy się z tym, co jest znane z obserwacji przypadków cholery. Zakażenie jest spowodowane przez spożycie pokarmu zakażonego dawką trucizny cholery, a każdy chory wydalą z organizmu zwiększoną dawkę tej trucizny, która następnie może skazić pokarm spożywany przez innych. Zatem charakterystyka mechanizmu przebiegu choroby w organizmie chorego wraz z opisem tego, jak trucizna cholery dostaje się do niego i jest przekazywana dalej, stanowi dla Snowa podstawę wyjaśnienia wszystkich zależności stwierdzonych dla zaobserwowanych zjawisk, np. wyższą zachorowalność w rejonach południowych miast, w pobliżu koryta rzeki, wśród zawodów związanych z obsługą statków.

Tak określony przez Snowa mechanizm przenoszenia i przebiegu cholery w organizmie został przez niego wyodrębniony na drodze rozważań teoretycznych, poprzez wyeliminowanie innych alternatywnych wyjaśnień symptomów cholery i jej obserwowanego przebiegu. Niezależnie jednak od tego, czy Snow podał słuszną hipotezę co do tego, jak przebiega ten mechanizm, każda inna musiałaby być sformułowana w kategoriach tych zjawisk, o których mówił Snow, a więc na poziomie oddziaływań aktywnego czynnika z komórkami organizmu pacjenta. Każde więc tego rodzaju wyjaśnienie mechanizmu przenoszenia i rozwoju cholery nie mogłoby spełnić empirystycznego wymogu adekwatności empirycznej, w związku z czym realność tej rzeczywistości empirysta umieściłby poza obszarem pytań epistemicznie relewantnych dla nauki. Kryterium adekwatności empirycznej byłoby jedynym kryterium rozstrzygającym między różnymi alternatywnymi wyjaśnieniami zaobserwowanych zależności, nieprzesądzającym jednak o ich realności. Empirysta zgodziłby się więc zaakceptować – chociaż odmówiłby

przekonania, że – opisany przez Snowa mechanizm, jako jedyny wówczas zaproponowany, który pozwalał wyjaśnić większość znanych zależności, zarówno dotyczących patologii tej choroby, jak i obserwacji geologicznych, społecznych itp. Nie miałyby jednak podstaw do tego, by zaakceptować, że to jest mechanizm *przyczynowy*.

Gdyby zaproponowany przez Snowa mechanizm przenoszenia i rozwoju cholery był jedyny, który – po wyeliminowaniu alternatywnych – byłby dopuszczalny teoretycznie, jako zgodny ze znaną patologią cholery, ale nie byłoby interakcji badawczej z tym mechanizmem, wówczas powiedzielibyśmy o nim, że jest przedmiotem akceptacji. Snow jednak przedsięwziął kroki konieczne do zaistnienia takiej interakcji. Ponieważ trucizna cholery była bezbarwna i bezzapachowa, więc Snow nie mógł precyzyjnie określić kryteriów bezpośredniej z nią interakcji. Takie kryteria ustalił jednak dla krwi zarażonych pacjentów (1936, 13-15). Interakcja badawcza Snowa z pacjentami chorymi na cholere, prowadząca do ustalenia, że zmiany w ich krwi są zgodne z innymi objawami cholery, oraz interakcja badawcza z innymi objawami choroby, prowadząca do ustalenia, że objawy specyficzne dla cholery są wcześniejsze od zmian w krwi i że brak jest ogólnych objawów zakażenia poprzedzających objawy cholery – to zasadniczy element przesądzający o istnieniu interakcji między badaczem a mechanizmem przenoszenia i rozwoju cholery. Podjęcie takiej interakcji stanowi także podstawę Snowa pełnego *przekonania* o istnieniu tego procesu.¹²

Jednak z tego, czego byłem w stanie dowiedzieć się o cholercie zarówno z moich własnych obserwacji, jak i opisów sporządzonych przez innych, wnioskuję, że cholera niezmiennie rozpoczyna się od zakażenia układu pokarmowego. (1936, 10)

Dla empiryzmu odróżnienie obserwowalne-nieobserwowalne, ze względu na kryterium adekwatności empirycznej, stało się jednym z najważniejszych dla nauki. Przykład Snowa to jeden z wielu, które pokazują, że przekonanie w nauce – jako związane z prawdziwością i znacznie mocniejsze epistemicznie od akceptacji – nie dotyczy tylko tego, co obserwowalne. W związku z tym można – bez zasadniczej szkody dla uprawiania nauk – dopuścić pewną niejasność w tym odróżnieniu, przypisując mu nie większe znaczenie, niż ma ono w faktycznie prowadzonych badaniach naukowych.

¹² Osobno można podjąć problem tego, czy dopuszczalna jest zmiana postawy epistemicznej z akceptacji na przekonanie ze względu na zaistnienie interakcji przyczynowej z badaną rzeczywistością.

Na gruncie metodologii proceduralnej można przyjąć, że przekonanie dotyczy tych zjawisk, z którymi można podjąć interakcję badawczą, rezerwując dla pozostałych nieprawdziwością postawę akceptacji. W ten sposób metodologia proceduralna wpisuje się w linię myśli obecną od początków w filozofii i nauce, że *o tym, co epistemicznie doniosłe, rozsądza kategoria przyczynowości* właśnie, a nie znacznie późniejsza pozytywistyczna idea obserwowalności.¹³

Pozostaje jednak pytanie, jak rozumieć stwierdzenie istnienia zależności przyczynowej, wzięwszy pod uwagę fakt, że – zgodnie z przyjętym tu podejściem kryterialnym – poza obszarem filozofii i metodologii nauk pozostawiam kwestię filozoficzną. Czy taki mechanizm przyczynowy faktycznie istnieje i co to znaczy, że on istnieje? Przede wszystkim więc kategoria przyczyny jest w metodologii proceduralnej pierwotna. Zakładam przy tym, że można określić kryteria zależności przyczynowych, traktując jako pierwotne pojęcie przyczyny w sensie interakcji badawczej, która jest elementem składowym każdego etapu badań naukowych, poczynając od organizowania kontekstu badań i zbierania danych, po ich analizę i uogólnienie.

O tym, że można wprowadzić do metodologicznej charakterystyki badań naukowych pojęcie przyczyny jako jednocześnie pierwotne i *nieredukowalne*, nie rozstrzygając przy tym filozoficznych pytań o naturę przyczynowości i pełną charakterystykę zależności przyczynowych, przekonuje m.in. możliwość skonstruowania takiej filozoficznej teorii przyczynowości, która nie przypisuje jej substancjalnego charakteru, ale jednocześnie broni jej realności. Sprawcza teoria przyczynowości, zainicjowana przez G. von Wrighta,¹⁴ a współcześnie kontynuowana przez P. Menziesia i H. Price'a, postuluje rozumienie zależności przyczynowej w perspektywie podmiotowego doświadczenia sprawstwa osobowego, przypisując jej ontologiczny status jakości wtórnej.¹⁵ Zgodnie z proponowaną przez Menziesia i Price'a

¹³ Podobną ideę współcześnie eksponuje stanowisko realizmu co do przedmiotów w obrębie szerszego nurtu tzw. nowego eksperymentalizmu (por. Hacking 1984; Giza 1990).

¹⁴ Por. też (Karpiński 1985, 97-99).

¹⁵ Zasadniczą ideę tej koncepcji G. von Wright wyraża zwięźle: „nie potrafimy zrozumieć przyczynowości ani odróżnienia między związkami nomicznymi a przypadkowymi jednorodnościami występującymi w naturze bez odwołania się do idei dotyczących robienia pewnych rzeczy oraz intencjonalnej interwencji w bieg natury” (von Wright 1971, 65-66). Konsekwencje takiego rozumienia przyczynowości dotyczą także innych podstawowych dla nauki kategorii metodologicznych, jak np. prawa naukowe: „konfirmacja praw nie jest jedynie kwestią powtarzania się udanych obserwacji. To jest sprawa «poddania prawa testowi». To, że taki test jest udany (z uwagi na prawdziwość prawa), oznacza, że uczymy się, jak robić jedne

wersją tej koncepcji, przyczynowość jest – jak kolor – pewnego rodzaju dyspozycją stwierdzaną w doświadczeniu wewnętrznym podmiotu, ale nadal realną. Podważa się tym samym pełny realizm, taki jak np. przyjęty przez Goldthorpe'a, a jednocześnie neguje redukcję przyczynowości do tego, co obserwowalne tylko, jak postuluje empiryzm konstruktywny.

Należy podkreślić, że projekt metodologii proceduralnej, który tu proponuję, nie wymaga uznania sprawczej teorii przyczynowości i może być uzgodniony z *inną* filozoficzną koncepcją, w szczególności taką, która przypisuje realność zależnościom przyczynowym. Sama jednak możliwość skonstruowania i obrony sprawczej teorii przyczynowości jest argumentem na rzecz tego, że przyczynowość jako pierwotna kategoria w metodologicznej charakterystyce badań naukowych nie musi być ani redukowalna do własności czysto obserwowalnych, ani nie jest w zasadniczy sposób związana ze stanowiskiem realizmu naukowego.

Zaproponowana przez van Fraassena charakterystyka nauki odwołuje się do celu nauki oraz do akceptacji teorii naukowych. W tym schemacie różnica między realizmem a empiryzmem sprowadzałaby się do tego, czy nauka ma na celu podawać teorie, które są prawdziwe czy empirycznie adekwatne, oraz do tego, czy akceptacja teorii związana jest z przekonaniem o jej prawdziwości czy empirycznej adekwatności (van Fraassen 1980, 9-12). Przyjęcie tezy realizmu, że teorie dają nam prawdziwy obraz świata, ukryty za zjawiskami, które można obserwować, daje nauce autorytet rozstrzygania tej kwestii niezależnie od tego, co jest dane w doświadczeniu. Zdaniem empirysty, jeśli ktoś twierdzi, że akceptuje teorię, bo jest przekonany o jej prawdziwości, to daje wyraz temu, że wierzy w jej prawdziwość pod względem tego, co daje się zaobserwować, *oraz* zaangażowaniu w tę teorię, które nie jest związane z prawdziwością, a daje wyraz postawy tej osoby jako gotowości do kontynuowania prac nad tą teorią, mimo jej niedoskonałości, zwłaszcza w zestawieniu z teatrami alternatywnymi.

Krytyka stanowiska realistycznego, jaką zwięźle tu przypomniałem za van Fraassenem, nie dotyczy metodologii proceduralnej. Przyjęcie kategorii przyczyny jako pierwotnej nie jest nieodłącznie związane, jak podkreślałem wcześniej, z uznaniem – na gruncie nauki, jej filozofii czy jej metodologii – przyczynowości za rzeczywistość leżącą *poza* zjawiskami. Taka teza jest

rzeczy, robiąc inne rzeczy (które już wiemy, jak robić), że nasze techniczne opanowanie natury zwiększa się. Można powiedzieć, że prawdziwości praw przyczynowych możemy być tak pewni, jak naszej zdolności robienia i sprawiania rzeczy". (von Wright 1971, 73)

oczywiście dopuszczalna, ale podkreślam – już nie na gruncie podejścia kryterialnego, lecz filozoficznego.

Argument antyrealisty – bez dodatkowych tez – jest w odniesieniu do innego niż nauka dyskursu całkowicie pozbawiony gruntu. Takim dodatkowym argumentem byłaby zasadnicza frontalna krytyka ontologii filozoficznej, co nie jest zadaniem łatwym ani jak dotąd – co na przykładzie licznych nieudanych prób pokazuje historia filozofii – skutecznym.

Bezpodstawny jest także zarzut kołowatości, jaki mogliby pod adresem metodologii proceduralnej sformułować realiści, jak Goldthorpe. Mogliby oni metodologii proceduralnej zarzucić, że przyjęcie kategorii przyczyny jako pierwotnej, a następnie określanie kryteriów odkrywania przyczyn, podane z odwołaniem się do niej, wprowadza kołowatość w rozważaniach.

W odpowiedzi podkreślam, że specyfiką podejścia proceduralnego jest wyodrębnienie filozoficznej kwestii *natury* czy *znaczenia* pojęcia przyczyny od kryterialnej kwestii *odkrycia* przyczyn oraz wyciągania z tego konsekwencji. Tę kwestię wyjaśniłem dokładniej w Rozdziale 9., analizując przykłady błędnego połączenia tych dwóch zadań. Po ich oddzieleniu nie ma podstaw do sformułowania zarzutu kołowatości. Mogą powstać co najwyżej trudności innej natury, a przede wszystkim zagadnienie spójności opracowanych metod odkrywania zależności przyczynowych, a w oparciu o nie – metod przewidywania skutków działań, z opracowaną niezależnie filozoficzną koncepcją przyczynowości. Tego rodzaju trudność nie wyznacza jakiejś *zasadniczej* bariery ani dla prowadzenia takich dociekań niezależnie od siebie, ani dla ich ewentualnego złączenia. Podobnie jak istnienie różnych interpretacji pojęcia prawdopodobieństwa, nie jest też powodem zaprzestania stosowania w badaniach naukowych np. metod statystycznych.

Ponadto, odkrycie przyczyn zjawisk – jak przyjmuje się w metodologii proceduralnej – dokonuje się zawsze poprzez przyczynową ingerencję, która ma na celu pozyskanie danych, aby następnie je przetworzyć na zależności między badanymi zmiennymi. Kołowatość w tym przypadku jest więc pozorna. Przy takim podejściu metodologia proceduralna jest bowiem na innym poziomie rozważań niż same modele przyczynowe czy sposoby tworzenia takich modeli. Konieczność wprowadzenia takich poziomów wynika przede wszystkim z tego, że poszczególne metody modelowania i odkrywania zależności przyczynowych wymagają określenia kryteriów wyższego rzędu niż sam poziom analizy, ze względu na które można oceniać trafność tych modeli. Nie dostrzegam lepszego sposobu wprowadzenia kategorii przyczyny do modeli zjawisk, zwłaszcza społecznych, niż proponowany w

metodologii proceduralnej. Przyjęcie kategorii przyczyny jako pierwotnej i nieredukowalnej nie przesądza jednoznacznie o jej filozoficznej interpretacji, ale nie jest też neutralne filozoficznie. Nie łączy się jednak z żadną filozoficzną koncepcją przyczynowości, pozostawiając to zadanie za obszarem filozofii i metodologii nauk.

To, że w badaniach naukowych nie należy rezygnować z kategorii przyczyny, jest, jak ilustruje przykład socjologów z Nuffield College w Oksfordzie czy uczonych w London School of Economics, zrozumiałe dla realistów. Zasadnicze zastrzeżenia zgłaszają przede wszystkim empiryści. Najnowsze prace nad modelowaniem zależności przyczynowych pokazują jednak (por. Rozdział 4. i Dodatek G), że nie można zrezygnować z kategorii przyczyny w modelowaniu zjawisk bez jednoczesnej rezygnacji z adekwatności empirycznej. Jak pokazuje przykład analizowany w Dodatku E, w niektórych modelach przyczynowych można wykazać istnienie obserwowalnych zależności między zmiennymi *tylko wtedy*, gdy przyjmie się istnienie zależności przyczynowych. W tym sensie dopuszczenie zależności przyczynowych musi być akceptowalne nawet dla empiryzmu konstruktywnego (van Fraassen 1980, 31).

Mimo dopuszczenia zależności przyczynowej wyłącznie ze względu na adekwatność empiryczną, empirysta nadal może kwestionować pierwotny charakter kategorii przyczyny w filozofii i metodologii nauk. Wracając więc do zarzutu kołowości, można by go kontynuować następująco. Nawet jeśli kryteria przyczyny formułuje się na innym poziomie niż same modele przyczynowe, to – na mocy postulatu metodologii proceduralnej – nie mamy do dyspozycji żadnej koncepcji przyczynowości ani wspierających ją i pozwalających ją egzemplifikować założeń ontologicznych. Takie kryteria pozwalają więc na wyodrębnienie tych modeli zależności między zmiennymi bądź tych podejść do modelowania przyczynowości, które spełniają podane kryteria. Nie mamy jednak żadnej gwarancji, że mają one jakikolwiek związek z rzeczywistością istniejącymi zależnościami przyczynowymi, gdyż zostały wyodrębnione w oparciu o intuicyjnie sformułowane kryteria. Konkluzją tego zarzutu zatem byłoby stwierdzenie, że nie da się obronić pierwotności i nieredukowalności przyczynowości bez przyjęcia pewnych rozstrzygnięć ontologicznych, choćby za cenę arbitralności tak opracowanych kryteriów.

Żadne ze stosowanych metod modelowania przyczynowości w badaniach nieeksperymentalnych nie dają gwarancji jednoznacznej poprawności wniosków w przypadku ogólnym, jak zostało to szczegółowo omówione w Rozdziale 4. Jeśli więc ten zarzut sprowadzałby się do przypisania subiektyw-

ności czy arbitralności, to przyjęcie założeń ontologicznych nie chroni przed nią lepiej niż np. wykorzystanie odpornych metod wnioskowania statystycznego. Ponadto można też poczynić taką uwagę, że powszechnie stosowane w epidemiologii Hilla jakościowe kryteria zależności przyczynowych (Zieliński i Mazurek 2002) nie odwołują się do procesu czy mechanizmu przyczynowego, lecz pozostają jedynie przy intuicyjnym rozumieniu tego terminu.

Przed wszystkim jednak odnotujmy, że ta trudność wydaje się oparta na zasadniczym nieporozumieniu. Metodologia proceduralna, jak podkreśliłem wcześniej, wymaga zasadniczego podziału zadań związanych z dociekaniami na temat przyczynowości. Trudność, na jaką być może zwraca się uwagę w tym zarzucie, dotyczy problemu zaznaczonego wcześniej – jak uzgodnić tak opracowane kryteria z wyartykułowaną filozoficznie teorią przyczynowości. To zadanie pozostawiamy poza obszarem niniejszej pracy, twierdząc, że jego rozwiązanie w niewielkim tylko stopniu mogłoby się przyczynić do doprecyzowania opracowanych tu kryteriów zależności przyczynowych i modelowania przyczynowego.¹⁶

Taka sama też jest odpowiedź dla tych, którzy oczekiwaliby pełniejszego określenia przyczyny niż tylko wprowadzenia jej jako kategorii pierwotnej i nieredukowalnej. Można tu także powtórzyć myśl, często przywoływaną w kontekście zarzutów stawianych modelom przyczynowym, które także nie eksplikują znaczenia tej kategorii – potrafimy odkrywać zależności przyczynowe wystarczająco dobrze, posługując się wyłącznie jej intuicyjnym rozumieniem. Oczywiście, nie jest wykluczone, że pełniejsza i ujednoznaczniająca charakterystyka filozoficzna przyczynowości pomogłaby w realizacji tego zadania. Wziąwszy jednak pod uwagę eksponencjalnie narastającą filozoficzną literaturę poświęconą temu tematowi, trudno byłoby oczekiwać pomocnych sugestii. Być może jest to kwestia przyszłości.

To, że z jednej strony istnieje inna alternatywa niż empiryzm i realizm naukowy, a z drugiej, że nie jest ona zauważana we współczesnych dyskusjach nad realizmem naukowym, wyjaśniałbym w taki sposób, iż – przywołując znane w literaturze odróżnienie Kazimierza Twardowskiego – dyskusje te skoncentrowane są na aspekcie wytworowym, a nie czynnościowym nauki. Tym samym metodologia proceduralna wpisuje się w tradycję szkoły

¹⁶ Niekiedy twierdzi się, że teoria kontrfaktycznej przyczynowości Lewisa wpłynęła znacząco na rozwój technik modelowania przyczynowego, zmienne kontrfaktyczne w podejściu RCM nie mają z tą koncepcją związku, a początek dały im prace J. Neymana, wcześniejsze o prawie cztery dekady.

Iwowsko-warszawskiej wyraźnego odróżniania czynnościowego ujmowania nauki od wytworowego i przypisania – zgodnie z programem metodologii pragmatycznej Kazimierza Ajdukiewicza (1975) – filozofii, a ogólnie humanistyce, kompetencji w zakresie przede wszystkim analiz czynnościowego aspektu nauki.

W konkluzji chciałbym zwrócić uwagę na perspektywy badawcze związane z pracami zostającymi poza głównym nurtem przywołanej wyżej tradycji¹⁷, jakie dostrzegam w niezwykle oryginalnej i zapoznanej dziś koncepcji przyczynowości É. Meyersona,¹⁸ kontynuowanej w Polsce przez Joachima Metallmanna (1934), a także w nieukończonych koncepcji przyczynowości Romana Ingardena (1945; 1955).¹⁹ Ich pełna dyskusja pozwoli, jak sądzę, opracować filozoficzną koncepcję przyczynowości, która byłaby dopełnieniem zaprezentowanego w tej książce stanowiska proceduralnego w kwestii kryteriów przyczynowości.

¹⁷ Apragmatyczność analiz pojęcia przyczynowości w szkole Iwowsko-warszawskiej podkreślają m.in. Woleński (1985, 228) oraz Bigaj (1994, 69).

¹⁸ Por. też (Sztejnburg 1931/1990, 20-31).

¹⁹ Analizę koncepcji Ingardena z wykorzystaniem niepublikowanych materiałów z jego wykładu o przyczynowości w UJ podaje Błaszczuk (1999).

Dodatki

A: Zmienne instrumentalne w modelu RCM

W artykule *Identification of causal effects using instrumental variables* J. Angrist i in. (1996) podejmują zagadnienie porównania modelu RCM z modelem SEM. Przedmiotem tego porównania są – wprowadzone w SEM – zmienne instrumentalne. Pokazują, że – przyjmując kilka prostych założeń – takie zmienne można wyrazić w modelu RCM. Dzięki zastosowaniu zmiennych instrumentalnych model RCM pozwala na wnioskowania przyczynowe w przypadku badań, w których przypisanie oddziaływania jest zaniedbywalne (szczególny przypadek to randomizacja), ale zgodność z podatnością na przypisane oddziaływanie nie jest pełna. Takie podejście – w przeciwieństwie do często stosowanej analizy „intencja-do-oddziaływania” (ITT), która ocenia efekty oddziaływania dla wszystkich członków grupy oddziaływania niezależnie od tego, czy faktycznie zostali poddani oddziaływaniu – pozwala na wprowadzenie opartego na zmiennych instrumentalnych oszacowania średniego efektu przyczynowego dla podgrupy jednostek, która jest poddana oddziaływaniu w sposób zgodny z mechanizmem oddziaływania. Mówiąc inaczej, wykorzystanie zmiennych instrumentalnych w modelu RCM pozwala na wnioskowanie przyczynowe w przypadku, gdy przypisanie oddziaływania jest losowe (zaniedbywalne), ale w poszczególnych grupach (kontrolnej i oddziaływania) nie ma pełnej zgodności z tym przypisaniem. Wnioski przyczynowe w tym przypadku estymują efekt przyczynowy

otrzymanego oddziaływania, a nie – jak w analizie ITT – efekt przypisania oddziaływania.¹

Przejdę teraz do szczegółowego wprowadzenia zmiennych instrumentalnych w SEM, a następnie w RCM, podając także za wspomnianymi autorami założenia wymagane przy przyczynowej interpretacji zmiennych instrumentalnych, a następnie analizę wrażliwości na pogwałcenie tych założeń w odniesieniu do opisywanego w omawianym tekście przykładu.

Zwróćmy przedtem uwagę, że modele SEM nie znalazły powszechnego zastosowania w pracach statystyków. Jako zasadniczy tego powód Angrist i in. podają dwie racje: wrażliwość tych modeli na kluczowe założenia (Little 1985) oraz ich niezdolność do odtwarzania wyników eksperymentalnych (Lalonde 1986). Ponadto podkreślają, że kluczowe założenia są wyrażane w kategoriach zakłóceń w niekompletnie określonych funkcjach regresji, zamiast w kategoriach zrozumiałych i potencjalnie obserwowalnych zmiennych:

Zwykle badacz nie ma jasnego pojęcia, co reprezentują te zakłócenia, i dlatego też trudno mu wyciągnąć realistyczne wnioski czy podać wyniki oparte na ich własnościach. (2004, 446)

Ich zdaniem, można natomiast w modelu RCM podać interpretację przyczynową opartego na zmiennej instrumentalnej oszacowania, pod warunkiem określenia kilku podstawowych założeń, omówionych poniżej.

Przykład dotyczy badania wpływu służby w wojsku na stan zdrowia, a dane dotyczą amerykańskich żołnierzy odbywających służbę w Wietnamie. Zmienna D reprezentująca oddziaływanie (odbywanie służby w wojsku lub nie) oraz Z reprezentująca przypisanie (status powołania do służby wojskowej, określany na podstawie numeru loterii – w przypadku wysokiego oznaczał brak powołania, w przypadku niskiego – powołanie do służby) są zmiennymi binarnymi. Gdyby zgodność z powołaniem była pełna, wówczas wszyscy z niskim numerem loterii ($Z = 1$) służyliby w wojsku ($D = 1$), a wszyscy z wysokimi numerami ($Z = 0$) nie służyliby ($D = 0$).

Zakłada się w tym przykładzie, że dla poszczególnych osób zaobserwowano wartości zmiennych Z , D oraz Y , która oznacza rezultat dla zdrowia danej osoby (też jako zmienna binarna). Reprezentując te wartości przez indeks dla danej osoby i , otrzymujemy następujący model równań strukturalnych wpływu służby w wojsku na zdrowie danej osoby (Angrist i in. 1996, 445):

¹ Por. (Frangakis i Rubin 2002; Mealli i Rubin 2002).

$$Y_i = \beta_0 + \beta_1 D_i + \varepsilon_i, \quad (\text{A.1})$$

$$D_i^* = \alpha_0 + \alpha_1 Z_i + v_i, \quad (\text{A.2})$$

gdzie $D_i = 1$, dla $D_i^* > 0$ oraz $D_i = 0$, dla $D_i^* \leq 0$.

W tym modelu β_1 reprezentuje efekt przyczynowy, jaki ma służba wojskowa D w stosunku do zdrowia wojskowych Y .

Dla identyfikacji współczynnika β_1 wprowadza się dwa założenia dotyczące zmiennej instrumentalnej Z . Pierwsze dotyczy tego, że Z_i nie jest skorelowana z błędami losowymi, zakłóceniami ε_i oraz v_i :

$$E[Z_i \varepsilon_i] = 0 \text{ oraz } E[Z_i v_i] = 0. \quad (\text{A.3})$$

Założenie, że korelacja między Z_i a ε_i jest zerowa oraz brak Z w Równaniu (A.1) są interpretowane w SEM jako fakt, że wpływ Z na Y jest całości wyrażony przez wpływ Z na D .

Drugie założenie dotyczy tego, że kowariancja między oddziaływaniem D_i a przypisaniem do danego oddziaływania Z_i jest niezerowa. Zmienna Z_i spełniająca te założenia nazywana jest zmienną instrumentalną w tym modelu. Oparty na niej estymator jest definiowany jako stosunek kowariancji w próbie:

$$\hat{\beta}_1^{IV} = \hat{\text{cov}}(Y_i, Z_i) / \hat{\text{cov}}(D_i, Z_i). \quad (\text{A.4})$$

W celu reprezentowania potencjalnych rezultatów RCM wymaga wprowadzenia tu pewnych modyfikacji w stosunku do SEM. I tak, \mathbf{Z} jest N -wymiarowym wektorem przypisań, którego i -tym elementem jest Z_i . $D_i(\mathbf{Z})$ jest wskaźnikiem tego, czy osoba i służyłaby w wojsku warunkowo względem losowo wygenerowanego wektora przypisań do powołania \mathbf{Z} . W przypadku pełnej podatności $D_i(\mathbf{Z}) = Z_i$ dla wszystkich i (każdy, kto miałby niski numer wylosowany na loterii powołań, służyłby w wojsku i nikt inny). Jest jednak wiele powodów, dla których taka zgodność nie zaistniała: mogą istnieć ochotnicy (którzy podejmą służbę niezależnie od wylosowanego numeru) oraz tacy, którzy ze względów rodzinnych czy zdrowotnych unikną służby (także niezależnie od wyniku losowania).

Podobnie modyfikuje się drugą ze zmiennych $Y_i(\mathbf{Z}, \mathbf{D})$, która reprezentuje odpowiedź osoby i warunkowo względem danego wektora wskaźników

służby wojskowej \mathbf{D} oraz wektora pierwszeństwa poboru do wojska \mathbf{Z} . Elementem i odpowiadającym odpowiedzi osoby i w N -elementowym wektorze $\mathbf{Y}(\mathbf{Z}, \mathbf{D})$ jest $Y_i(\mathbf{Z}, \mathbf{D})$. $D_i(\mathbf{Z})$ oraz $Y_i(\mathbf{Z}, \mathbf{D})$ reprezentują potencjalne rezultaty, które są stałe, lecz nieznanne, częściowo tylko zaobserwowane poprzez przypisanie oddziaływania poszczególnym jednostkom. Różnice w tych potencjalnych wynikach wynikają z przypisanych i faktycznie otrzymanych oddziaływań, co ujawnia analiza danych otrzymanych przez losowe przypisanie \mathbf{Z} w skończonej populacji N jednostek badania.

Pierwsze z przyjętych przez Angrist i in. założeń dotyczy braku interferencji między poszczególnymi jednostkami badań. Za Rubinem (1978, 1980, 1990b) określa ją jako SUTVA (ang. Stable Unit Treatment Value Assumption):

- (i) jeśli $Z_i = Z'_i$, wtedy $D_i(\mathbf{Z}) = D_i(\mathbf{Z}')$;
 (ii) jeśli $Z_i = Z'_i$ oraz $D_i = D'_i$, wtedy $Y_i(\mathbf{Z}, \mathbf{D}) = Y_i(\mathbf{Z}', \mathbf{D}')$.

Założenie SUTVA oznacza, że potencjalny rezultat dla każdej osoby i jest niezwiązany z oddziaływaniem na inne osoby, a więc można pominąć oddziaływania na inne osoby: $D_i(\mathbf{Z}) = D_i(Z_i)$ oraz $Y_i(\mathbf{Z}, \mathbf{D}) = Y_i(Z_i, D_i)$.

Zgodnie z zaproponowaną przez Rubina (1974) definicją efektu przyczynowego otrzymujemy w tym przypadku definicję:

$$\text{efekt przyczynowy } Z \text{ dla jednostki } i \text{ na } D \text{ to: } D_i(1) - D_i(0) \quad (\text{A.5})$$

$$\text{oraz efekt przyczynowy } Z \text{ na } Y \text{ to: } Y_i(1, D_i(1)) - Y_i(0, D_i(0)).$$

Pierwszy z tych efektów oznacza różnicę rezultatu wyciągnięcia losu z niskim numerem (przypisanie do grupy odbywającej służbę wojskową) dla faktycznego odbycia służby wojskowej w stosunku do rezultatu wylosowania losu z wysokim numerem (przypisanie do grupy nieodbywającej służby wojskowej) dla faktycznego nieodbycia służby wojskowej.

Drugi z tych efektów to różnica rezultatu dla zdrowia danej jednostki, która została przypisana do odbycia służby wojskowej i jej odpowiedzi na to przypisanie, w stosunku do rezultatu dla zdrowia tej samej jednostki, gdyby nie została przypisana i odpowiedziała na to przypisanie.

Ponieważ odpowiedź jednostki na przypisanie do grupy oddziaływania lub kontrolnej (odpowiednio $D_i(1)$ i $D_i(0)$) może być niezgodna z przypisaniem (np. ochotnicy czy zwolnieni ze służby), więc mamy do czynienia z efektami przyczynowymi w sytuacji niepełnej zgodności. Dlatego też An-

grist i in. używają tu określenia stosowanego w próbach medycznych, a mianowicie „efekty leczenia zamierzonego” (ang. intention-to-treat effects, w skrócie ITT).²

Kolejne założenie w ogólnej postaci dotyczy zaniedbywalności przypisania oddziaływania (Rubin 1978; Rosenbaum i Rubin 1983). Jednak dla uproszczenia dyskusji Angrist i in. (1996, 446) wprowadzają je w postaci randomizacji przypisania danego rodzaju oddziaływania Z_i :

$$\Pr(\mathbf{Z} = \mathbf{c}) = \Pr(\mathbf{Z} = \mathbf{c}') \quad (\text{A.6})$$

dla wszystkich \mathbf{c} oraz \mathbf{c}' takich, że $\iota^T \mathbf{c}$ oraz $\iota^T \mathbf{c}'$, gdzie ι jest N -wymiarowym wektorem kolumnowym dla wszystkich elementów równych jeden.

Założenia SUTVA oraz losowego przypisania pozwalają na wyprowadzenie nieobciążonych estymatorów średnich efektów ITT. Otrzymuje się je, za Neymanem (1923/1990), jako różnicę między średnią Y oraz D z próby, gdzie obie te zmienne są sklasyfikowane zgodnie z wartościami Z (a więc jako różnice średnich między grupą oddziaływania a grupą kontrolną).

Nieobciążony estymator średniego efektu przyczynowego Z na Y to:

$$\frac{\sum_i Y_i Z_i}{\sum_i Z_i} - \frac{\sum_i Y_i (1 - Z_i)}{\sum_i (1 - Z_i)} \quad (\text{A.7})$$

Z kolei nieobciążony estymator średniego efektu przyczynowego Z na D ma postać:

$$\frac{\sum_i D_i Z_i}{\sum_i Z_i} - \frac{\sum_i D_i (1 - Z_i)}{\sum_i (1 - Z_i)} \quad (\text{A.8})$$

Stosunek tych estymatorów jest równy estymatorowi opartemu na zmiennej instrumentalnej (stosunek kowariancji zmiennych Y_i i Z_i do kowariancji D_i i Z_i), określonego w Równaniu (A.4). Na tej podstawie uznaje się równość opartego na zmiennej instrumentalnej oszacowania (granica opar-

² Efekt typu ITT mierzy się dla wszystkich jednostek w grupie oddziaływania (wszystkich, którzy zostali przypisani do grupy oddziaływania) niezależnie od tego, czy faktycznie wypełniali program zgodny z zaleceniami dla tego oddziaływania, np. pacjent, który nie przyjmuje zaleconych leków, czy taki, który wycofał się ze studium.

tego na niej estymatora) ze stosunkiem średnich efektów ITT (Angrist i in., 1996, 447).

Jeśli jednak ustala się średni efekt przyczynowy, jaki ma odbycie służby wojskowej D na stan zdrowia żołnierzy Y , to różnica w średnich rezultatach dla grup o różnym oddziaływaniu nie dostarcza nieobciążonego czy nawet niespójnego estymatora. W celu zdefiniowania takiego estymatora dla D i Y wprowadza się dalsze założenia.

Założenie wykluczenia wymaga, aby przypisanie oddziaływania nie było związane z potencjalnym rezultatem, jeśli uwzględnione jest już samo oddziaływanie:

$$Y(\mathbf{Z}, \mathbf{D}) = Y(\mathbf{Z}', \mathbf{D}) \text{ dla wszystkich } \mathbf{Z} \text{ oraz } \mathbf{Z}' \quad (\text{A.9})$$

i dla wszystkich \mathbf{D} .

Z tego założenia wynika, że stan zdrowia danej osoby Y_i zależy tylko od tego, czy faktycznie odbyła służbę wojskową, niezależnie od przypisania oddziaływania: $Y_i(1, d) = Y_i(1, d)$ dla $d = 0, 1$. Podobnie więc jak w modelu SEM, wpływ, jaki wylosowanie numeru związanego z loterią powołań Z ma na zdrowie żołnierza, jest w całości zapośredniczony przez to, czy ta osoba faktycznie odbyła służbę wojskową D .

Założenie wykluczenia, na co należy tutaj zwrócić uwagę, zestawia dwie wielkości, których nie można jednocześnie zaobserwować: albo dana osoba miała los o niskim numerze i otrzymała powołanie do wojska, albo o wysokim numerze i nie otrzymała takiego powołania. To założenie nie jest więc bezpośrednio weryfikowalne na podstawie danych, jak przyznają Angrist i in. (1996, 447), lecz ma sprawdzalne konsekwencje, wyprowadzalne w oparciu o dwa pierwsze założenia (SUTVA oraz założenie randomizacji).

Założenie wykluczenia pozwala na następujące zdefiniowanie potencjalnego rezultatu $Y(\mathbf{Z}, \mathbf{D})$:

$$Y(\mathbf{Z}, \mathbf{D}) = Y(\mathbf{Z}', \mathbf{D}) = Y(\mathbf{D}) \text{ dla wszystkich } \mathbf{Z} \text{ oraz } \mathbf{Z}' \quad (\text{A.10})$$

i dla wszystkich \mathbf{D} .

Na tej podstawie oraz na mocy założenia SUTVA $Y_i(\mathbf{Z}, \mathbf{D}) = Y_i(D_i)$. Efekt przyczynowy D w stosunku do Y wyraża się następująco:

efektem przyczynowym D w stosunku do Y (A.11)

dla danej osoby i jest $Y_i(1) - Y_i(0)$.

Oba człony też różnicy nie są obserwowalne dla tej samej osoby, dlatego Angrist i in. proponują skoncentrowanie się na średnim efekcie przyczynowym. Średnią dla dowolnej funkcji $g(-)$ zdefiniowanej dla danej populacji oznaczmy jako $E[g]$, a średnią tejże funkcji dla podpopulacji wyznaczonej przez pewną wartość h_0 innej funkcji $h(-)$ jako $E[g | h(-) = h_0]$, gdzie rozmiar tej podpopulacji oznaczany jest jako $P[h(-) = h_0] = E[1_{h(-) = h_0}]$.³

Kolejne z założeń, jakie wprowadzają Angrist i in. dotyczy niezerowego efektu Z w stosunku do średniego prawdopodobieństwa oddziaływania, czyli D . Zgodnie z nim ten efekt $E[D_i(1) - D_i(0)]$ jest różny od zera.

Piąte założenie stwierdza, że nie istnieje taka osoba, która zachowywałaby się dokładnie odwrotnie, niż wynikałoby to z przypisania oddziaływania Z (po wylosowaniu niskiego numeru unikałaby służby wojskowej, a po wylosowaniu wysokiego – starałaby się o podjęcie służby). To założenie określane jest jako monotoniczność: $D_i(1) \geq D_i(0)$ dla wszystkich jednostek w populacji (dla każdego $i = 1, \dots, N$).

Połączenie założeń czwartego i piątego, z których wynika, że zachodzi ostra nierówność $D_i(1) > D_i(0)$ dla co najmniej jednej osoby w populacji, określane jest jako mocna monotoniczność.

Założenia powyżej omówione, a więc założenie SUTVA, losowe przypisanie, ograniczenie wykluczenia, niezerowy efekt średni Z w stosunku do D oraz monotoniczność, prowadzą do definicji zmiennej instrumentalnej na gruncie modelu RCM (Angrist i in. 1996, 447-48):

Zmienna Z jest *instrumentalna* w stosunku do efektu przyczynowego D w stosunku do Y , jeśli Z ma niezerowy efekt w stosunku do D , spełnia założenie ograniczenia wykluczenia i monotoniczności, jest przypisana losowo (zaniechwalnie) oraz zachodzi SUTVA.

Założenia SUTVA oraz ograniczenie wykluczenia pozwalają z kolei na ustalenie zależności, jaka zachodzi między efektem ITT Z w stosunku do Y oraz D , a efektem przyczynowym, jaki ma D w stosunku do Y :

$$Y_i(1, D_i(1)) - Y_i(0, D_i(0)) = (Y_i(1) - Y_i(0)) (D_i(1) - D_i(0)). \quad (\text{A.12})$$

³ $1_{\{-\}}$ to funkcja wskaźnikowa.

Efekt przyczynowy, jaki ma Z w stosunku do Y dla danej osoby i , wyraża się więc iloczynem efektu przyczynowego Z w stosunku do D oraz D w stosunku do Y .

Zapisując średni efekt przyczynowy Z w stosunku do Y jako sumę ważoną średnich efektów dla dwóch podpopulacji $D_i(1) \neq D_i(0)$ oraz stosując założenie o monotoniczności (zgodnie z nim różnica $D_i(1) - D_i(0)$ jest równa 1 lub 0), otrzymujemy (Angrist i in. 1996, 448):

$$E[Y_i(1, D_i(1)) - Y_i(0, D_i(0))] = \quad (\text{A.13})$$

$$E[Y_i(1) - Y_i(0) | D_i(1) - D_i(0) = 1] \cdot P[D_i(1) - D_i(0) = 1]$$

Otrzymujemy więc twierdzenie, które ustala związek między oszacowaniem opartym na zmiennej instrumentalnej a efektem przyczynowym D w stosunku do Y :

$$\frac{E[Y_i(1, D_i(1)) - Y_i(0, D_i(0))]}{E[D_i(1) - D_i(0)]} = \quad (\text{A.14})$$

$$= E[Y_i(1) - Y_i(0) | D_i(1) - D_i(0) = 1].$$

Efekt przyczynowy wyrażony powyższym wzorem Angrist i in. określają jako średni lokalny efekt oddziaływania (ang. Local Average Treatment Effect, w skrócie LATE).

Interpretacja tego efektu jest następująca. Dla jednostek działających zgodnie z przypisanym oddziaływaniem (podejmują służbę wojskową po wylosowaniu niskiego numeru) – podatni, dla których $D_i(1) - D_i(0) = 1$, efekt przyczynowy Z w stosunku do Y wynosi $Y_i(1) - Y_i(0)$. Dla tych, którzy unikają służby wojskowej niezależnie od wylosowanej liczby w loterii powołań do wojska (ang. zawsze niechętni, tu: unikający poboru, czyli osoby dla których $D_i(1) = D_i(0) = 0$), oraz dla tych, którzy podejmą służbę niezależnie od wyniku loterii (ang. zawsze chętni, tu: wolontariusze – dla których $D_i(1) = D_i(0) = 1$), a więc ci wszyscy, dla których $D_i(1) - D_i(0) = 0$, efekt przyczynowy Z w stosunku do Y jest zerowy (na mocy założenia o ograniczeniu wykluczenia). Ostatnia grupa to tzw. buntownicy, czyli osoby działające dokładnie przeciwnie do przypisanego oddziaływania, dla których $D_i(1) - D_i(0) = -1$. W ich przypadku efekt przyczynowy Z w stosunku do Y wynosi $Y_i(0) - Y_i(1)$. Te trzy ostatnie grupy: zawsze chętni, zawsze niechętni

oraz buntownicy stanowią grupę osób niepodatnych, czyli działających niezgodnie z przypisanym oddziaływaniem.

Twierdzenie A.14, wyprowadzone w oparciu o przyjęte założenia, prowadzi do wniosku, że średni efekt przyczynowy Z w stosunku do Y jest proporcjonalny do średniego efektu przyczynowego D w stosunku do Y dla osób działających zgodnie z przypisanym oddziaływaniem.⁴

Po przedstawieniu obu sposobów reprezentowania (SEM oraz RCM) zależności przyczynowych między zmiennymi Z , D oraz Y przejdę do ich zestawienia. Angrist i in. zwracają uwagę na niejasność założeń prowadzących w SEM do stwierdzenia, że Z jest niezależne od zakłóceń w równaniach tego modelu (A.1) oraz (A.2). W modelu RCM to założenie jest wyeksplikowane jako dwa różne założenia dotyczące potencjalnie obserwowalnych własności. Pierwsze z nich to założenie o zrandomizowanym (zaniebdywalnym) przypisaniu oddziaływania (w ogóle nie dotyczy bezpośrednich efektów przypisania). Drugie założenie dotyczy ograniczenia wykluczenia (w ogóle nie dotyczy mechanizmu przypisania). Prowadzi ono do uwzględnienia przez badacza efektu określonego rodzaju oddziaływania na rezultat. Pogwałcenie tych dwóch założeń może mieć różne źródła i różne konsekwencje, dlatego też Angrist i in. uznają, że

połączenie tych dwóch założeń w jedno, dotyczące zerowej korelacji między zmiennymi instrumentalnymi a zakłóceniami, prowadzi do nieporozumień odnoszących się do istoty założeń związanych z identyfikowalnością, a ponadto utrudnia ocenę i komunikację wiarygodności leżącego u podstaw modelu. (1996, 450)

Komentując twierdzenie A.14, zwróciłem uwagę, że założenie ograniczenia wykluczenia pozwala wyróżnić cztery podstawowe podpopulacje i ocenić możliwe pogwałcenia tego założenia (a w przypadku, gdyby można było dokonać identyfikacji grupy podatnych i niepodatnych, to założenie można bezpośrednio sprawdzić). W przypadku wolontariuszy, którzy podjęliby służbę wojskową niezależnie od wyniku loterii powołań, można przyjąć, że zmienna instrumentalna Z nie ma wpływu na ich zdrowie (ten efekt jest w całości zapośredniczony przez D , a w tym przypadku wiadomo, że niezależnie od przypisania oddziaływania służba wojskowa zostanie podjęta).

W przypadku tych, którzy uniknęliby służby wojskowej niezależnie od wyniku loterii, można rozważyć dwa przypadki. Dla tych, którzy są zwol-

⁴ Angrist i in. omawiają także kwestie dotyczące estymacji efektu przyczynowego oraz identyfikowalności poszczególnych podpopulacji (1996, 449).

nieni ze względu na stan zdrowia, przypisanie Z pozostanie bez wpływu na ich stan zdrowia Y . Jednak w przypadku tych, którzy pozostali w szkołach lub wyjechali za granicę, by uniknąć służby, można wskazać wpływ przypisania na bieg ich życia, a tym samym na stan zdrowia, co byłoby także pogwałceniem założenia ograniczenia wykluczenia (Angrist i in. 1996, 449).

Druga składowa przyjętego w SEM założenia o niezależności Z od zakłóceń to wyłożone tu założenie o randomizowanym (zaniedbywalnym) przypisaniu. W omawianym przykładzie to założenie jest spełnione, gdyż powołania do służby wojskowej dokonywały się poprzez losowy mechanizm generujący poszczególne liczby w loterii powołań. Angrist i in. podkreślają jednak, że mimo iż randomizacja zmiennej losowej jest wystarczająca do nieobciążonej estymacji średniego efektu oddziaływania Z w stosunku do Y oraz średniego efektu oddziaływania Z w stosunku do D , to nie wynika z tego, że oszacowanie oparte na zmiennej losowej może być zinterpretowane jako średni efekt przyczynowy (1996, 450).

Kolejne z założeń w podejściu RCM, jakim jest monotoniczność wykluczająca istnienie buntowników, nie ma odpowiednika w SEM. Jest jednak używane *implicite* przy stosowaniu stałych parametrów dla relacji między Z_i a D_i . Założenie o stałości parametrów jest jednak zbyt mocne, chociaż postulowanie niezerowej kowariancji między oddziaływaniem a przypisaniem nie jest wystarczające do zinterpretowania opartego na zmiennej instrumentalnej oszacowania jako średniego efektu przyczynowego.

Kolejna uwaga dotyczy tego, że w modelu SEM nie ma odróżnienia między efektem dla całej populacji a efektem dla podpopulacji podatnych. LATE (średni lokalny efekt przyczynowy) jest, zdaniem Angrista i in., strukturalny w sensie reprezentowania związku przyczynowego, ale nie jest strukturalny w sensie reprezentowania parametru niezmienniczego dla różnych populacji.⁵

Podsumowaniem omawianego przykładu jest zastosowanie użycia zmiennych instrumentalnych w modelu RCM do sformułowania opartego na nich i zinterpretowanego przyczynowo oszacowania. Potencjalnym rezultatem w tym badaniu jest zmienna wskaźnikowa $Y_i(z, d)$, której wartość wynosi 1 dla osób zmarłych pomiędzy rokiem 1974 a 1983 i mających wynik loterii z oraz wskaźnik stanu służby wojskowej równy d . Śmiertelność w tym

⁵ Kolejną zaletą podejścia RCM ma być to, że można zastosować analizę wrażliwości na odchylenia od poszczególnych założeń, w diskutowanym przypadku dotyczy to przede wszystkim założenia ograniczenia wykluczania oraz monotoniczności (Angrist i in. 1996, 450-51).

czasie niezwiązana ze służbą wojskową Y_i to śmiertelność cywilna. W analizowanym przykładzie założenie SUTVA jest interpretowane jako brak związku między odbyciem służby przez daną osobę (oraz stanem jej zdrowia) a wynikiem losowania dla innych osób w wieku poborowym. Założenie dotyczące losowości mechanizmu przypisania jest spełnione dla loterii powołań. Ograniczenie wykluczenia oznacza tu, że stan powołania do służby wojskowej nie ma związku z ryzykiem śmiertelności cywilnej, pod warunkiem uwzględnienia stanu służby wojskowej. Założenie o niezerowym średnim efekcie Z w stosunku do D oznacza, że wylosowanie niskiego numeru w loterii powołań zwiększa prawdopodobieństwo odbycia służby wojskowej. Założenie monotoniczności z kolei prowadzi do stwierdzenia, że nie ma takiej osoby, która w razie wylosowania niskiego numeru unikałaby służby, a w przypadku wylosowania wysokiego – dążyłaby do odbycia służby wojskowej (Angrist i in. 1996, 452).

Analiza dostępnych danych z wykorzystaniem oszacowania opartego na zmiennej instrumentalnej prowadzi omawianych autorów do wniosku, że służba wojskowa wpływała na stan zdrowia. I tak, w przypadku jednej z analizowanych podpopulacji, jaką był rocznik 1950, wpływ przyczynowy służby wojskowej na ryzyko śmiertelności zwiększał prawdopodobieństwo śmiertelności o ok. 25%.

B: S. Wrighta model przyczynowy SEM

W 1925 r. Departament do Spraw Rolnictwa Stanów Zjednoczonych opublikował biuletyn *Corn and hog correlations* autorstwa Wrighta, w którym podane zostały empiryczne estymatory dla wielorównaniowego modelu. Uwzględniono w nim cenę zboża y_1 , cenę tuczników latem y_2 i zimą y_3 , ich płodność y_4 , średnią wagę latem y_5 oraz ich ilość zimą y_6 . Do wyjaśnienia tych zjawisk Wright wykorzystał następujące zmienne: plony zbóż x_1 , opóźnioną cenę zbóż x_2 , i podwójnie opóźnioną ich cenę x_3 , opóźnioną płodność tuczników x_4 , oraz podwójnie x_5 i potrójnie x_6 opóźnioną, a także opóźnioną cenę zimą x_7 . W tym modelu ε reprezentuje zakłócenia, a zmienne y_4 oraz x_4 , x_5 i x_6 są niezaobserwowane. W efekcie otrzymujemy równania dla wszystkich zmiennych wyjaśnianych:

$$y_1 = a_1x_1 + \varepsilon_1, \quad (\text{B.1})$$

$$y_2 = gy_1 + hx_2 + ix_4 + jx_5 + \varepsilon_2,$$

$$y_3 = ky_1 + lx_2 + mx_4 + \varepsilon_3,$$

$$y_4 = by_2 + ay_1 + cx_2 + ex_3 + fx_4 + dx_7 + \varepsilon_4,$$

$$y_5 = e_1y_4 + e_2x_4 + e_3x_5 + \varepsilon_5,$$

$$y_6 = f_1y_4 + f_2x_4 + f_3x_5 + f_4x_6 + \varepsilon_6.$$

Korzystając z zerowych korelacji między obserwowalnymi zmiennymi, podał on oparte na zmiennych instrumentalnych estymatory parametrów strukturalnych (określanych przez Wrighta jako współczynniki ścieżek), przy czym te estymatory miały w możliwie najdokładniejszy sposób spełniać możliwie najliczniejszą grupę równań.

Wpływ Wrighta na rozwój modeli strukturalnych był zapośredniczony przez H. Schultza (Goldberger 1972, 985-86). W swojej pracy *The theory and measurement of demand* (1938) zasadniczy problem charakteryzował następująco:

Czy możliwe jest statystyczne wydedukowanie teoretycznej krzywej popytu (lub podaży), gdy znane są tylko [...] punkty [ich] przecięcia [...] w różnych momentach czasu? (1938, 72-80; cyt. za: Goldberger 1972, 980)

Odpowiedź pozytywna na to pytanie w przypadku krzywej popytu polega na odniesieniu cen p do ilości produktu q w danym roku, a dla krzywej podaży – przez odniesienie cen p_{-1} do ilości produktu q w następnym roku. Schultz przyjął więc następujący model pajęczynowy:

$$q^D = ap + u, \quad q^S = \beta p_{-1} + v, \quad q^D = q^S \quad (\text{B.2})$$

Ten model jest identyfikowalny jedynie dzięki opóźnionej cenie p_{-1} . Dla estymowania β należy przeprowadzić w tym modelu regresję q na p_{-1} , a dla estymowania $1/\alpha$ regresję p na q . Jak w komentarzu podkreśla Goldberger (1972, 980-81), Schultz – ani żaden z jego następców w ekonometrii – nie wykorzystał w pełni znanej sobie teorii Wrighta. Dlatego też model Schultza (B.2), mimo uwzględnienia zagadnień identyfikowalności oraz estymacji, jest bardzo ograniczony: zawiera tylko dwa równania, jedyną podstawą identyfikacji jest opóźniona cena i całkowicie został pominięty związek między specyfikacją modelu a estymacją.

W publikacji z 1925 r. Wright tylko wspomniał możliwość wprowadzenia funkcji popytu w modelu zależności zboża i tuczników (B.1), gdyż to wymagałoby poszerzenia analizy ścieżkowej. W apendyksie do książki jego ojca P. G. Wrighta (1929), przyjmują oni statyczny model popytu-podaży:

$$q^D = ap + u, \quad q^S = \beta p + v, \quad q^D = q^S. \quad (\text{B.3})$$

Następnie przyjmują zmienną instrumentalną z (cena substytutu lub indeks dobrobytu), o której zakładają, że nie jest skorelowana z zakłóceniem (ang. disturbance) krzywej podaży: $Cov(z, v) = 0$. Z tego faktu oraz z modelu (B.3) wynika, że $\beta = Cov(z, q) / Cov(z, p)$. Podobnie dla $\alpha = Cov(x, q) / Cov(x, p)$, gdzie x to plony na akr (lub opóźniona cena).

Model (B.3) w zredukowanej postaci:

$$q = (-\beta u + \alpha v) / (\alpha - \beta), \quad p = (-u + v) / (\alpha - \beta) \quad (\text{B.4})$$

dał Wrightowi podstawę (1934) do stwierdzenia, że trzy estymowalne wielkości, a mianowicie $V(q)$, $V(p)$ oraz $C(q, p)$ są niewystarczające do zi-

dentyfikowania pięciu nieznanymi parametrów modelu: α , β , $V(u)$, $V(v)$ oraz $C(u, v)$.

Jedno z rozwiązań polega na wprowadzeniu zmiennej instrumentalnej z , która reprezentuje czynnik przesuwający krzywą popytu, ale nie podaży:

$$u = \gamma z + u', \quad \text{gdzie } C(v, z) = 0. \quad (\text{B.5})$$

W zredukowanej postaci otrzymujemy więc zmodyfikowaną postać modelu (B.3) ojca i syna Wrightów:

$$q = (-\beta\gamma z - \beta u' + \alpha v) / (\alpha - \beta), \quad p = (-\gamma z - u' + v) / (\alpha - \beta), \quad (\text{B.6})$$

w którym jest pięć estymowalnych wielkości: $V(q)$, $V(p)$, $C(q, p)$, $C(z, q)$ oraz $C(z, p)$ i sześć parametrów: α , β , γ , $V(u')$, $V(v)$ oraz $C(u', v)$.

Przy założeniu, że korelacja między krzywą popytu i podaży jest zerowa, a więc, że $C(u', v) = 0$, możliwe jest rozwiązanie, jakie proponuje Wright (1934, 199). Gdyby jednak to założenie było niespełnione, wówczas można wprowadzić inną zmienną instrumentalną x , która reprezentowałaby czynnik przesuwający krzywą podaży, ale nie popytu:

$$v = \delta x + v', \quad \text{gdzie } C(x, u) = 0. \quad (\text{B.7})$$

W rezultacie otrzymujemy siedem estymowalnych wielkości dla siedmiu parametrów.¹

Tę metodę estymacji Wright wykorzystał następnie w tym samym artykule do estymacji modelu popytu-podaży dla tuczników.

Komentując późniejszą pracę Schultza, który udostępnił niektóre dane Wrightowi, a także wielokrotnie dyskutował z nim na temat modelowania przyczynowego, Goldberger podkreśla, że „Wydawałoby się czymś rozsądnym oczekiwać, że Schultz, pisząc w 1938 roku, będzie przywoływał prace Wrighta.” Tak się jednak nie stało, a przywołania są jedynie drugorzędne. Goldberger ubolewa na tym faktem:

Trudno wyjaśnić, dlaczego Schultz nie dostrzegwał, ani nie umiał wykorzystać, bardzo silnego podejścia Wrighta do formułowania i estymacji modeli równań strukturalnych. (Goldberger 1972, 985)

¹ Te wielkości oraz parametry to odpowiednio: $V(q)$, $V(p)$, $C(q, p)$, $C(z, q)$, $C(z, p)$, $C(x, q)$, $C(x, p)$ oraz α , β , γ , δ , $V(u')$, $V(v')$ oraz $C(u', v')$.

Wright podał mu następujące wyjaśnienie:

Henry był tak oddany w pełni formalnemu i obiektywnemu podejściu w analizie wielozmiennowej, że nigdy nie udało mi się nakłonić go do dostrzeżenia możliwej użyteczności mojego nieco nieformalnego i subiektywnego sposobu podejścia. (rozmowa, wrzesień 1971; cyt. za: Goldberger 1972, 986)

Podobny brak zainteresowania osiągnięciami Wrighta Goldberger dostrzega u następców Schultza, jak H. Working, T. Koopmans, P. Fisk czy H. Wold (1972, 986-87).

C: Autonomiczność relacji ekonomicznych – T. Haavelma model konsumpcji

Pojęcie autonomiczności relacji zilustruję, posługując się przykładem podanym przez Haavelma (1944, 31-33) i przyjmując jego notację. Niech x oznacza przypadający na głowę stosunek konsumpcji danego produktu w grupie ludzi o takich samych dochodach R , p – cenę produktu, P – indeks kosztów życia. O następującej funkcji popytu założmy, że jest prawdziwa:

$$x = a \frac{p}{P} + b \frac{R}{P} + c + \varepsilon, \quad (\text{C.1})$$

gdzie a , b oraz c są stałymi, a ε jest zmienną losową o małej wariancji. Warunkowa wartość oczekiwana x jest więc dana wzorem:

$$E(x | \frac{p}{P}, \frac{R}{P}) = a \frac{p}{P} + b \frac{R}{P} + c. \quad (\text{C.2})$$

O Równaniu (C.1) zakłada się, że jest autonomiczne w tym sensie, że dla dowolnych wartości $\frac{p}{P}$ oraz $\frac{R}{P}$ wartość x może być estymowana przez (C.2).¹

Niech p_0 , P_0 oraz R_0 oznaczają średnie wartości odpowiednich zmiennych takie, że w porównaniu z nimi wartości $(p - p_0)$, $(P - P_0)$ oraz $(R - R_0)$ są małe. Otrzymujemy więc przybliżoną wartość równania (C.1):

$$x = a \frac{p_0 + (p - p_0)}{P_0 + (P - P_0)} + b \frac{R_0 + (R - R_0)}{P_0 + (P - P_0)} + c + \varepsilon. \quad (\text{C.3})$$

Po przekształceniach otrzymujemy równanie:

$$x = a \left(\frac{p + p_0}{P_0} \right) + b \left(\frac{R + R_0}{P_0} \right) - P \left(a \frac{p_0}{P_0^2} + b \frac{R_0}{P_0^2} \right) + c + \varepsilon', \quad (\text{C.4})$$

¹ Parametry a oraz b są więc w tym sensie strukturalne, a więc nie zależą od danego rozkładu wartości zmiennych p , P oraz R (Pearl 2000, 27-28).

gdzie ε' to nowy termin resztowy, uwzględniający dodatkowo błędy powyższej aproksymacji.²

Może zdarzyć się taka sytuacja, że zaobserwowane wartości zmiennych p , P oraz R są funkcją czasu. Wówczas dla tych zaobserwowanych danych istnieje inna relacja o takiej samej formie, jak (C.4), ale o innych współczynnikach, która będzie jeszcze lepiej dopasowana do danych. Taka jednak relacja jest konfluentna, a więc nie jest aproksymacją funkcji popytu (C.1) ani nie pozwala na przewidywanie efektu dowolnej zmiany ceny lub dochodu.

Kolejny ważny przykład podany przez Haavelma (1944, 33-38) dotyczy nie tyle zagadnienia autonomiczności zależności w odniesieniu do estymacji i przewidywania, lecz – uważanego dziś za drugi istotny problem wnioskowania przyczynowego – zmian polityki ekonomicznej, a więc takiego działania, które modyfikuje niektóre z dotychczasowych zależności przyczynowych. Omawiany przykład dotyczy K. Wicksella teorii stóp procentowych i cen produktów. Upraszczające założenia w tym przypadku są następujące: w danej społeczności są tylko jednostki, firmy prywatne oraz banki. Jednostki dzielą dochód na konsumpcję (lub zwiększenie zasobów gotówki) i na oszczędzanie na lokatach terminowych w bankach (jedyne źródło oszczędności w tej społeczności). Firmy są jedynym producentem dóbr, kierującym się wyłącznie oczekiwanym zyskiem, a inwestycje podejmują tylko za pośrednictwem kredytów bankowych. Cały dochód rozdzielają między jednostki. Ceny dóbr i usług zmieniają się w czasie i są reprezentowane przez wspólną zmienną – poziom cen. Istnieje tylko jedna stopa procentowa. Otrzymujemy więc następujące zmienne: $S(t)$ – wszystkie oszczędności przypadające na jednostkę czasu, $I(t)$ – wartość wszystkich inwestycji w jednostce czasu, $\rho(t)$ – bankowa stopa procentowa w punkcie t , $P(t)$ – poziom cen w punkcie t oraz $R(t)$ – całkowity dochód danej społeczności w jednostce czasu.

Zasadnicze pytanie dotyczy wpływu polityki kredytowej banków na ceny. W uproszczonym modelu, który ma być podstawą do udzielenia odpowiedzi, zakłada się następujące zależności. Podaż na oszczędności wyraża się funkcją:

$$S(t) = a_0 + a_1 \rho(t) + a_2 P(t) + a_3 \dot{P}(t) + a_4 R(t). \quad (\text{C.5})$$

² Przy małej wariancji zmiennych ε' jest praktycznie nieodróżnialne od ε .

Określa ona, że podaż oszczędności w postaci lokat bankowych zależy pozytywnie od stopy procentowej i całkowitego dochodu, a negatywnie od poziomu cen oraz stopnia zmiany poziomu cen $\dot{P}(t)$ – pochodna poziomu cen reprezentująca oczekiwaną przyszłą rzeczywistą wartość zaoszczędzonych pieniędzy.

Z kolei popyt na kredyty bankowe jest zależny negatywnie od stopy procentowej i pozytywnie od poziomu cen, a w sposób teoretycznie nieokreślony³ od stopnia zmiany poziomu cen:

$$I(t) = b_0 + b_1 \rho(t) + b_2 P(t) + b_3 \dot{P}(t). \quad (\text{C.6})$$

Z kolei Haavelmo proponuje założenie, że banki działają bezinteresownie:

$$I(t) = S(t). \quad (\text{C.7})$$

Z powyższych równań (C.5), (C.6) oraz (C.7) wynika, że każdej z wartości zmiennych $R(t)$, $P(t)$ oraz $\dot{P}(t)$ odpowiada oparta na równowadze rynkowej stopa procentowa $\bar{\rho}(t)$, określana przez Wicksella jako *normalna* stopa:

$$\bar{\rho}(t) = A_0 + A_1 P(t) + A_2 \dot{P}(t) + A_3 R(t), \quad (\text{C.8})$$

gdzie A_i reprezentują współczynniki: $A_0 = (b_0 - a_0) / (a_1 - b_1)$, $A_1 = (b_2 - a_2) / (a_1 - b_1)$, $A_2 = (b_3 - a_3) / (a_1 - b_1)$ oraz $A_3 = a_4 / (a_1 - b_1)$.

Jeśli banki chcą wpłynąć na zmianę ilości pieniędzy pozostających poza nimi, muszą ustalić stopę procentową $\rho(t)$ na poziomie odpowiednio różnym od (C.8).

Na podstawie równań podaży oszczędności (C.5) oraz popytu na kredyty (C.6) otrzymujemy równanie określające inflację:

$$I(t) - S(t) = (b_1 - a_1) [\rho(t) - \bar{\rho}(t)]. \quad (\text{C.9})$$

³ Haavelmo zauważa w tym miejscu, że kiedy wzrasta poziom cen, firmy mogą oczekiwać, że kupią środki produkcji na tańszym rynku niż ten, na którym później sprzedadzą swoje produkty, co może być zachętą do inwestowania (1944, 35).

Wynika z niego, że inflacja jest negatywnie proporcjonalna w stosunku do różnicy między aktualną bankową stopą procentową a stopą normalną, określoną w (C.8).

Przyjmuje się, że inflacji – którą Haavelmo określa jako „barometr wydawanych pieniędzy” – towarzyszy proporcjonalny wzrost cen:

$$\dot{P}(t) = k [I(t) - S(t)]. \quad (\text{C.10})$$

Z (C.9) oraz (C.10) wynika uproszczona wersja twierdzenia Wicksella, dotyczącego wpływu bankowych stóp procentowych (odbiegających od normalnej) na ceny:

$$\dot{P}(t) = k (b_1 - a_1) [\rho(t) - \bar{\rho}(t)]. \quad (\text{C.11})$$

Zasadnicza kwestia dotyczy tego, czy zależność określająca inflację (C.9), jej wpływ na wzrost cen (C.10), oraz zależność wzrostu cen od bankowej stopy procentowej (C.11) są autonomiczne. W tym kontekście pytanie to dotyczy przede wszystkim niezmienniczości tych zależności w przypadku zmiany bankowej polityki kształtowania stóp procentowych.

W odniesieniu do zależności określającej inflację (C.9) zauważmy za Haavelmem, iż została ona wyprowadzona z równania podaży oszczędności (C.5) oraz popytu na kredyty (C.6), bez nakładania jakichkolwiek ograniczeń na funkcję określającą bankową stopę procentową $\rho(t)$, poziom cen $P(t)$ czy całkowity dochód $R(t)$. Ponadto z równania określającego normalną stopę procentową (C.8) wynika, że dla dowolnego poziomu cen i dochodu całkowitego istnieje normalna stopa procentowa $\bar{\rho}(t)$, podczas gdy dla zadanego momentu czasu normalnej stopie procentowej $\bar{\rho}(t)$ odpowiada nieskończenie wiele funkcji czasu dla poziomu cen i dochodu całkowitego. Haavelmo konkluduje więc, że zależność określająca inflację (C.9) jest autonomiczna – pozwala bowiem dla arbitralnie wybranych funkcji opisujących bankową stopę procentową $\rho(t)$ oraz normalną stopę procentową $\bar{\rho}(t)$ w zależności od czasu obliczyć inflację.

Ta własność zależności wyrażanej w Równaniu (C.9), jak podkreśla Haavelmo,

nie jest matematyczną własnością tego równania – nie można jej odkryć, patrząc na nie. Opiera się ona na hipotezie dotyczącej tego, jak różnica $I(t) - S(t)$ zachowywałaby się *faktycznie* dla różnych arbitralnych zmian stopy procentowej i normalnej stopy procentowej. (1944, 36)

Takiej własności nie ma równanie – zaproponowane w hipotetycznym modelu alternatywnym do omawianego tu – o tej samej postaci, ale w którym zakłada się istnienie liniowych trendów w czasie. I tak, $I(t) - S(t) = mt$, a $\rho(t) - \bar{\rho}(t) = nt$. To alternatywne w stosunku do (C.9) równanie ma postać:

$$I(t) - S(t) = \frac{m}{n} [\rho(t) - \bar{\rho}(t)]. \quad (\text{C.12})$$

(C.12) jako alternatywne równanie inflacji w stosunku do (C.9) nie jest autonomiczne, gdyż przy takiej polityce kształtującej stopy procentowe, która wprowadziłaby niezgodną z trendem $\rho(t) - \bar{\rho}(t) = nt$ zmianę, stałoby się fałszywe.

Zależność wzrostu cen od inflacji (C.10) Haavelmo uznaje także za autonomiczną. Niezależnie bowiem, jaka jest inflacja, pozwala na obliczenie tempa zmiany poziomu cen. W mniejszym stopniu natomiast autonomiczność przysługuje zależności tempa zmiany poziomu cen od zmiany stóp procentowych (C.11). Klasa funkcji czasu, które spełniają to równanie, jest bowiem klasą tych funkcji, które spełniają jednocześnie Równanie (C.9) i (C.10).

Dyskutowany dotąd uproszczony model Wicksella można uzupełnić o założenie dotyczące polityki kredytowej banków. Dotąd bowiem zmienna $\rho(t)$ funkcjonowała jako wolny parametr. Można jednak przyjąć, że banki przez pewien okres czasu realizują określoną politykę względem stóp procentowych. Haavelmo proponuje następującą charakterystykę tej polityki w pewnym okresie czasu. Kiedy banki zdają sobie sprawę, że inflacja jest pozytywna, podnoszą stopę procentową, aby zabezpieczyć swoją płynność finansową, i na odwrót – obniżają stopę procentową, gdy inflacja jest negatywna:

$$\dot{\rho}(t) = c [I(t) - S(t)], \quad (\text{C.13})$$

gdzie c jest stałą dodatnią. Z zależności określającej inflację (C.9) wynika, że

$$\dot{\rho}(t) = c (b_1 - a_1) [\rho(t) - \bar{\rho}(t)]. \quad (\text{C.14})$$

Na podstawie uproszczonej wersji twierdzenia Wicksella (C.11) otrzymujemy zaś

$$\dot{P}(t) = \frac{k}{c} \rho(t). \quad (\text{C.15})$$

Zgodnie z tą zależnością poziom cen zmienia się w tym samym kierunku, co stopy procentowe. Postawmy za Haavelmem pytanie, czy ta zależność jest autonomiczna? A więc czy pozwala na obliczenie przypuszczalnego – lub jakbyśmy powiedzieli w dzisiejszej terminologii „kontraktycznego” – wpływu dowolnej polityki kształtującej stopy procentowe na poziom cen? Odpowiedź na to pytanie jest negatywna. Zależność (C.15) zachodzi tylko w tym przypadku, gdy dochód całkowity $R(t)$, wielkość inwestycji $I(t)$, całkowite oszczędności $S(t)$, poziom cen $P(t)$, stopa procentowa $\rho(t)$ oraz normalna stopa procentowa $\bar{\rho}(t)$ spełniają dokładnie takie funkcje czasu, jakie określone są w zależnościach (C.8), (C.9), (C.10) oraz (C.13). Zmiana którejkolwiek z tych zależności prowadziłaby do podważenia zależności (C.15), a tym samym ukazywałoby się, że nie jest ona niezmiennicza względem niektórych rodzajów podejmowanej polityki ekonomicznej.

Autonomiczną zależność, która pozwala na przewidywanie efektów, Haavelmo wyprowadza na podstawie zależności określającej normalną stopę procentową (C.8) oraz na podstawie uproszczonej wersji twierdzenia Wicksella (C.11):

$$\dot{P}(t) + BP(t) = H_1 \rho(t) + H_2 R(t) + H_0, \quad (\text{C.16})$$

gdzie B , H_1 , H_2 oraz H_0 są stałymi, zależnymi od występujących w (C.8) i (C.11). Ta zależność nie wprowadza ograniczeń co do funkcji $\rho(t) + R(t)$, ponieważ pozwala dla dowolnych wartości stopy procentowej oraz dochodu całkowitego obliczyć poziom cen $P(t)$ jako ich funkcję.

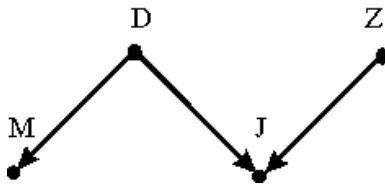
W konkluzji na pytanie, skąd wiadomo, że należy korzystać z Równania (C.16), a nie (C.15), Haavelmo podkreśla, że

Nie ma formalnej metody, za pomocą której można dojść do takiego wniosku. W rzeczy samej, zaczynając od innego modelu z innymi założeniami, można dojść do przeciwnego wniosku. Aby podjąć decyzję, musimy *wiedzieć* lub wyobrazić sobie – na podstawie ogólnego doświadczenia – która z dwóch zależności, (C.15) czy (C.16), byłaby *rzeczywiście* bardziej stabilna, jeśliby wykorzystała jedną z nich jako relację autonomiczną. (1944, 38)

D: Przykład uaktualniania w modelu przyczynowym BN

Mechanizm uaktualniania rozkładu prawdopodobieństwa w sieciach bayesowskich zilustruję prostym przykładem.¹ Załóżmy, że Jan mieszkający pod Warszawą zaobserwował pewnego ranka, że trawa przed jego domem jest mokra. Mogło więc padać ($D = d$) lub zapomniał wczoraj wieczorem wyłączyć zraszacza ($Z = t$). Po tym spostrzeżeniu przekonanie o tym, że padał deszcz oraz o tym, że nie wyłączył zraszacza, wzrasta na sile. Jan rozgląda się następnie wokół i dostrzega, że trawnik sąsiadującej z nim Marii jest także mokry, więc jest niemal pewien, że padał deszcz. Tym samym siła przekonania o tym, że zostawił wczoraj wieczorem włączony zraszacza, powraca do stanu wyjściowego.²

Wykorzystując dostępne informacje, a więc także wiedzę ogólną, można skonstruować graf przedstawiony na Rysunku D.1.



Rysunek D.1.

Strzałki na tym grafie interpretowane są jako zależności przyczynowe, np. na to, czy trawnik Marii jest mokry ($M = m$), wpływa tylko to, czy padał deszcz ($D = d$), a nie ma tu znaczenia to, czy Jan włączył swój zraszacza ($Z = z$). Można więc, upraszczając, przyjąć, że podstawowe informacje, jakie są w tym przypadku dostępne, określają jednoczesny stan tych czterech zjawisk względem siebie dla poszczególnych obserwacji. Można więc na tej podstawie obliczyć, jak często trawnik Marii jest mokry, gdy nie pada deszcz, a Jan zostawił na noc włączony zraszacza.

¹ Por. też (Lad 1999).

² W tym przykładzie dla lepszej prezentacji dość swobodnie posługuję się różnymi interpretacjami pojęcia prawdopodobieństwa, można jednak tę samą analizę przeprowadzić bez zmian, zachowując jednorodność przy ujmowaniu prawdopodobieństwa.

Zapisując systematycznie te obserwacje (i częstości zjawisk), można byłoby podać łączny rozkład prawdopodobieństwa dla uwzględnionych tu zmiennych. To żmudne zadanie można jednak uprościć, przyjmując założenie Markowa (por. Punkt 4.4). Z własności Markowa dla grafu przedstawionego na rysunku D.1 wynika, że rozkład łączny przybiera postać:

$$P(JMDZ) = P(J|MDZ) \cdot P(M|DZ) \cdot P(D|Z) \cdot P(Z) \quad (D.1)$$

Uwzględniając niezależności, które są zilustrowane na Rysunku D.1 jako brak krawędzi (między J a M , M a Z oraz D a Z), można uprościć rozkład łączny do postaci:

$$P(JMDZ) = P(J|DZ) \cdot P(M|D) \cdot P(D) \cdot P(Z) \quad (D.2)$$

Założmy ponadto – przyjmując dla uproszczenia rachunków binarność wszystkich występujących tu zmiennych – że w odniesieniu do analizowanego przykładu posiadamy informacje o rozkładach, np. częstościach wystąpień poszczególnych zdarzeń, przedstawionych w Tabeli D.1.

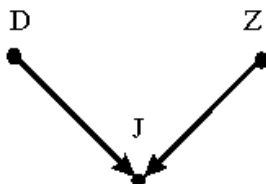
Lp.	Rozkład/model	Rozkład zajścia	Rozkład niezajścia
1.	$P(D)$	$P(d) = 0,2$	$P(\bar{d}) = 0,8$
2.	$P(Z)$	$P(z) = 0,1$	$P(\bar{z}) = 0,9$
3.	$P(M D)$	$P(m d) = 1$	$P(\bar{m} d) = 0$
4.		$P(m \bar{d}) = 0,2$	$P(\bar{m} \bar{d}) = 0,8$
5.	$P(J DZ)$	$P(j dz) = 1$	$P(\bar{j} dz) = 0$
6.		$P(j \bar{d}z) = 0,9$	$P(\bar{j} \bar{d}z) = 0,1$
7.		$P(j d\bar{z}) = 1$	$P(\bar{j} d\bar{z}) = 0$
8.		$P(j \bar{d}\bar{z}) = 0$	$P(\bar{j} \bar{d}\bar{z}) = 1$

Tabela D.1.

Dane te można zinterpretować następująco. W porze roku, której dotyczy analizowany przykład, deszcz pada ok. 6 razy w miesiącu, natomiast Janowi zdarza się zostawić włączony zraszacz 3 razy w miesiącu (jeśli oczywiście używa go w tej porze roku codziennie). Zawsze, ilekroć pada deszcz, trawnik Marii jest mokry, natomiast zdarza się mniej więcej 6 razy na miesiąc, że jest mokry nawet wtedy, gdy deszcz nie pada. W przypadku trawnika Jana wiadomo z całą pewnością, że jeśli pada deszcz, to trawnik jest mokry, natomiast gdy deszcz nie pada, a włączony jest zraszacz, to na ogół trawnik

jest mokry – tylko raz na dziesięć włączeń zraszacza jest suchy. Wiadomo także z całą pewnością, że nie ma innej możliwości, żeby trawnik Jana był mokry, niż przez deszcz czy działania zraszacza.

Kolejny krok polega na wywnioskowaniu siły przekonania Jana o tym, że jego trawnik jest mokry, *zanim* Jan obejrzał swoją trawę rano. Z Rysunku D.1 możemy odczytać, że na to przekonanie wpływ mają częstość występowania deszczu i włączania zraszacza. Dla obliczenia stopnia przekonania $P(J)$ posłużę się zatem podmodelem, uwzględniając tylko istotne czynniki:



Rysunek D.2.

Dla podmodelu $P(J|D Z)$ rozkład łączny oblicza się, korzystając z (nie)zależności reprezentowanych na grafie przedstawionym na Rysunku D.2:

$$P(J D Z) = P(J|D Z) \cdot P(D) \cdot P(Z) \quad (D.3)$$

Korzystając z Równania (D.3) oraz niezależności poszczególnych zdarzeń, zgodnie z grafem na Rysunku D.2, w oparciu o dane zawarte w Tabeli D.1, oblicza się rozkład łączny $P(J D Z)$.

Lp.	Rozkład łączny	Rozkład zajścia
1.	$P(J D Z)$	$P(j d z) = 1 \cdot 0,2 \cdot 0,1 = 0,02$
2.		$P(\bar{j} d z) = 0$
3.		$P(j \bar{d} z) = 0,9 \cdot 0,8 \cdot 0,1 = 0,072$
4.		$P(\bar{j} \bar{d} z) = 0,1 \cdot 0,8 \cdot 0,1 = 0,008$
5.		$P(j d \bar{z}) = 1 \cdot 0,2 \cdot 0,9 = 0,18$
6.		$P(\bar{j} d \bar{z}) = 0$
7.		$P(j \bar{d} \bar{z}) = 0$
8.		$P(\bar{j} \bar{d} \bar{z}) = 1 \cdot 0,8 \cdot 0,9 = 0,72$

Tabela D.2.

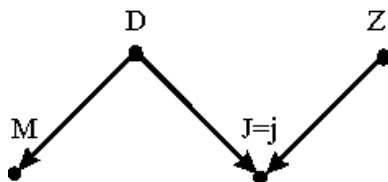
Z rozkładu dla tego podmodelu wylicza się stopień przekonania Jana o tym, że jego trawnik jest mokry $P(j)$, marginalizując D oraz Z z rozkładu łącznego $P(J D Z)$. Ponieważ poszczególne zdarzenia są niezależne, więc:

$$P(j) = P(j d z) + P(j \bar{d} z) + P(j d \bar{z}) + P(j \bar{d} \bar{z}) = 0,272. \quad (D.4)$$

O wiele silniejsze jest przekonanie Jana o tym, że trawnik nie jest mokry:

$$P(\bar{j}) = 0,728. \quad (D.5)$$

Dotychczas uwzględniane były tylko informacje znane Janowi przed spojrzeniem rano na trawę. Po uzyskaniu tej informacji ($J = j$) zmienia się jego dotychczasowy stopień przekonania.



Rysunek D.3.

Tę zmianę określa się, uwzględniając model przyczynowy (por. Rysunek D.13) oraz wnioskowanie w tym modelu, przebiegające zgodnie z twierdzeniem Bayesa:

$$P(J D Z | j) = \frac{P(J D Z) \cdot P(j | J D Z)}{P(j)} \quad (D.6)$$

Oczywiście w przypadku, gdy trawa jest sucha ($J = \bar{j}$), to $P(j | J D Z) = 0$, natomiast gdy jest mokra ($J = j$), to $P(j | J D Z) = 1$. Formuła D.6 upraszcza się zatem do postaci:

$$P(J D Z | j) = \frac{P(J D Z)}{P(j)}, \quad (D.7)$$

gdzie $J = j$ lub równa jest zeru w przeciwnym przypadku.

Zgodnie z bayesowską regułą kondycjonalizacji przekonanie po uzyskaniu nowych informacji jest aktualizowane i staje się przekonaniem wyjściowym przed kolejną obserwacją lub eksperymentem:

$$P(JD Z \bar{j}) = P_1(JD Z). \quad (\text{D.8})$$

Wykorzystujemy dane z poprzednich obliczeń dla określenia aktualnego stanu przekonań Jana (por. Tabela D.3).

Z tego modelu po marginalizacji Z oraz J otrzymujemy:

$$P_1(d) = P_1(j d z) + P_1(j d \bar{z}) = 0,736 \quad (\text{D.9})$$

Podobnie po marginalizacji D oraz J otrzymujemy:

$$P_1(z) = P_1(j d z) + P_1(j \bar{d} z) = 0,338 \quad (\text{D.10})$$

Lp.	Rozkład łączny	Rozkład zajścia
1.	$P_1(JDZ)$	$P_1(j d z) = 0,2/0,272 = 0,074$
2.		$P_1(j \bar{d} z) = 0,072/0,272 = 0,264$
3.		$P_1(j d \bar{z}) = 0,181/0,272 = 0,662$
4.		$P_1(j \bar{d} \bar{z}) = 0$
5.		0 w pozostałych przypadkach (gdy $J = \bar{j}$).

Tabela D.3.

Uaktualnienione dzięki nowym informacjom wartości $P_1(d)$ oraz $P_1(z)$ reprezentują nową siłę przekonań Jana o tym, że padał deszcz oraz że został włączony zraszacz – to pierwsze jest ponad dwa razy mocniejsze niż drugie i prawie czterokrotnie silniejsze niż przed dokonaniem obserwacji.

Poranne spostrzeżenie Jana zmienia też jego przekonanie co do trawnika jego sąsiadki. W tym przypadku uwzględniamy podmodel grafu z Rysunku D.2, który dotyczy trawnika Marii, i tego, jaki wpływ ma na niego deszcz.



Rysunek D.4.

Obliczamy rozkład łączny tego podmodelu (uwzględniając niezależność M oraz $J = j$):

$$P_1(MD) = P_1(M|D) \cdot P_1(D) = P(M|D, j) \cdot P_1(D) = \quad (D.11)$$

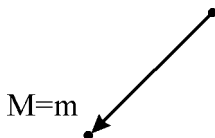
$$P(M|D) \cdot P_1(D) = \frac{P(M|D)}{P(D)} \cdot P(D) \cdot P_1(D) = P(MD) \cdot \frac{P_1(D)}{P(D)}$$

Na tej podstawie obliczamy rozkład dla podmodelu (por. Tabela D.4).

Lp.	Rozkład łączny	Rozkład zajścia
1.	$P_1(MD)$	$P_1(md) = 0,2 \cdot (0,736/0,2) = 0,736$
2.		$P_1(m\bar{d}) = 0,16 \cdot (0,264/0,8) = 0,0528$
3.		$P_1(\bar{m}d) = 0$
4.		$P_1(\bar{m}\bar{d}) = 0,64 \cdot (0,264/0,8) = 0,2112$

Tabela D.4.

Z opisu, który podałem na początku analizy tego przykładu, wiadomo, iż Jan dokonał kolejnego spostrzeżenia, a mianowicie, że trawnik Marii też jest mokry ($M = m$).



Rysunek D.5.

Ta obserwacja, zgodnie z twierdzeniem Bayesa, prowadzi do kolejnego uaktualnienia stopni przekonań Jana:

$$P_1(MD | m) = P_2(MD) = P(MD | \bar{m}). \quad (D.12)$$

Zastosowanie do dyskutowanego podmodelu twierdzenia Bayesa daje:

$$P_1(MD | m) = \frac{P_1(MD)}{P_1(m)} \cdot P(m | MD). \quad (D.13)$$

Podobnie jak w przypadku poprzedniej obserwacji, dla przypadków, gdy $M = \bar{m}$, $P(m | MD) = 0$, a więc i $P_1(MD | m) = 0$. Natomiast gdy $M = m$, wówczas $P(m | MD) = 1$:

$$P_1(MD | m) = \frac{P_1(MD)}{P_1(m)}. \quad (D.14)$$

Dla analizowanego podmodelu otrzymujemy zatem uaktualniony rozkład:

Lp.	Rozkład łączny	Rozkład zajścia
1.	$P_2(MD)$	$P_2(m d) = 0,736/0,7888$
2.		$P_2(m \bar{d}) = 0,0528/0,7888$
3.		$P_2(\bar{m} d) = 0$
4.		$P_2(\bar{m} \bar{d}) = 0$

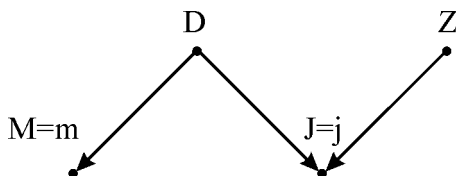
Tabela D.5.

A następnie, marginalizując M w $P_2(MD)$, otrzymujemy siłę przekonania Jana o tym, że padał deszcz, po uwzględnieniu obu obserwacji:

$$P_2(d) = P_2(m d) + P_2(\bar{m} d) = 0,93. \quad (D.15)$$

Jan jest więc niemal pewien tego, że padał deszcz, a siła jego przekonania, w porównaniu do sytuacji poprzedzającej dokonanie obserwacji, wzrosła ponad 4,5 raza.

A co z siłą przekonania o alternatywnej przyczynie tego, że trawnik Jana był rano mokry? Wracamy zatem do podmodelu $P(JDZ)$.



Rysunek D.6.

Stosując twierdzenie Bayesa oraz wykorzystując niezależności widoczne na grafie na Rysunku D.6, obliczamy rozkład łączny dla tego modelu, uwzględniając obserwacje, jakich dokonał Jan:

$$\begin{aligned}
 P_2(M D Z) &= P_1(M D Z | m) = & (D.16) \\
 &= P_1(J | D Z m) \cdot P_1(D | Z m) \cdot P_1(Z | m) \\
 &= P_1(J | D Z) \cdot P_1(D | m) \cdot P_1(Z) = P_1(J | D Z) \cdot (P_1(D)/P_1(D)) \\
 &\cdot P_1(Z) \cdot P_1(D | m) = P_1(M D Z) \cdot P_2(D) \cdot P_1(D).
 \end{aligned}$$

Marginalizując J oraz D w $P_2(M D Z) = P_1(M D Z) \cdot P_2(D) \cdot P_1(D)$, otrzymujemy:

$$P_2(z) = P_1(j d z) \cdot \frac{P_2(d)}{P_1(d)} + P_1(j \bar{d} z) \cdot \frac{P_2(\bar{d})}{P_1(\bar{d})} = 0,163. \quad (D.17)$$

Przypomnę, że wyjściowe przekonanie Jana o tym, że zapomniał wyłączyć zraszacz, gdy przestał go już używać, wynosi $P(m | \bar{d}) = 0,2$, a więc, mimo że Jan zaobserwował, że trawa i jego, i sąsiadki jest mokra i wszystko wskazuje, że to właśnie jest skutkiem deszczu, to jednak pozostał cień wątpliwości, że Jan mógł też zostawić na noc włączony zraszacz. Tak też można tłumaczyć niewielki wzrost przekonania Jana o tym, że zraszacz został włączony na noc $P_2(z) - P(z) = 0,063$.

E: Przykład modelu przyczynowego BN ze zmiennymi ukrytymi

Ilustracji takich problemów dostarcza klasyczny już myślowy przykład eksperymentu (Cochran 1957, 266-67; Wainer 1989, 128-29), w którym fumiganty glebowe (F) wykorzystuje się do kontrolowania populacji węgorka (W), co ma wpłynąć na wzrost plonów z uprawy owsa (P). Fumiganty mogą wpływać (zarówno pozytywnie, jak negatywnie) na plony upraw owsa bezpośrednio, niezależnie od zmian populacji węgorka. Zasadniczy problem dotyczy w tym przypadku określenia całkowitego wpływu fumigantów na plony upraw (a więc bezpośrednio i pośrednio – przez regulację populacji węgorka), przy uwzględnieniu komplikacji, jakie są spowodowane wystąpieniem innych czynników.

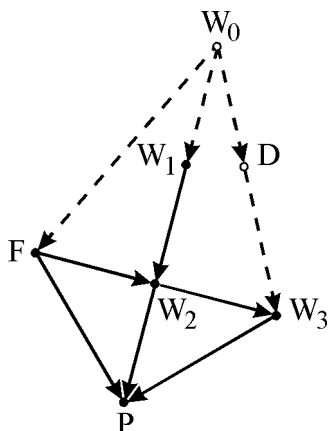
Najprostszy sposób polegałby na przeprowadzeniu kontrolowanego eksperymentu. To jednak często w warunkach naturalnych jest niewykonalne, gdyż rolnicy sami chcą decydować, które pola zostaną poddane fumigacji. Ponadto decyzję o ilości zastosowanych fumigantów rolnicy podejmują, biorąc pod uwagę wielkość populacji węgorka w roku poprzednim (F_0), a to jest nie tylko wielkością nieznaną, ale i mocno skorelowaną z tegoroczną populacją.

Można jednak wykorzystać fakt, że analiza laboratoryjna próbek gleby pozwala ustalić populację węgorka przed (W_1) i po (W_2) zastosowaniu fumigantów. Ponadto, ponieważ fumiganty są aktywne tylko przez krótki okres, więc można bezpiecznie założyć, że nie mają wpływu na wzrost tej populacji węgorka (W_3), która przetrwa fumigację. Ten wzrost zależy więc od populacji ptaków i innych drapieżników (D), co jest skorelowane z zeszłoroczną populacją węgorka, a więc i samą fumigacją.

W oparciu o te rozważania można, posługując się procedurą opisaną przez Pearla (2000, s. 49-54) oraz Spirtesa, Glymoura i Scheinesa (2000, s. 73-155), skonstruować diagram przyczynowy,¹ podany na Rysunku E.1.

Wielkości mierzalne, takie jak populacja węgorka przed (W_1) i po fumigacji (W_2) oraz po zakończeniu sezonu (W_3), a także plony upraw owsa (P) i ilość fumigantów (F), reprezentowane są przez kropki. Wielkości nieznanne (lub niemierzalne), jak W_0 oraz populacja ptaków (i innych drapieżników) D , są reprezentowane przez małe okręgi.

¹ Szczegółowe omówienie konstrukcji różnych, też nieeksperymentalnych, typów modeli podaje Brzeziński (1996, 343-430).



Rysunek E.1. (Opracowano na podstawie: Pearl 1995).

Zmienne na diagramie połączone dwoma rodzajami krawędzi: liniami ciągłymi w przypadku zmiennych mierzalnych i liniami przerywanymi w przypadku zmiennych nieznanych (lub niemierzalnych). W obu przypadkach strzałka sygnalizuje kierunek zależności przyczynowej. Zgodnie z założeniem Markowa (por. Punkt 4.4) każda ze strzałek reprezentuje autonomiczny, tj. całkowicie niezależny od pozostałych, mechanizm przyczynowy. Natomiast brak połączeń między zmiennymi oznacza negatywne założenie przyczynowe o nieistnieniu między nimi związku przyczynowego. Na przykład brak krawędzi między W_1 a P oznacza, że populacja węgorka nie wpływa na plony z upraw przed fumigacją, a cały wpływ jest zapośredniczony przez tę populację po fumigacji, W_2 oraz W_3 .

Wykorzystując wypracowane przez Pearla reguły rachunku „zrób” (2000, 85-89), można określić całkowity wpływ fumigantów na plony upraw:

$$p(p \parallel f^*) = \sum_{w_1} \sum_{w_2} \sum_{w_3} p(p \mid w_2, w_3, f) \cdot p(w_2 \mid w_1, f) \sum_{f^*} p(w_3 \mid w_1, w_2, f^*) p(w_1, f^*), \quad (\text{E.1})$$

gdzie $p(p \parallel f^*)$ oznacza prawdopodobieństwo osiągnięcia plonów na poziomie $P = p$, pod warunkiem, że fumigacja jest ustalona (przez niezależną interwencję) na poziomie $F = f^*$.

Zgodnie z własnością Markowa rozkład prawdopodobieństwa dla reprezentowanych w grafie zmiennych odzwierciedla niektóre własności struktury

przyczynowej, przede wszystkim zerowe korelacje cząstkowe między zmiennymi. Graficznym tego odpowiednikiem jest brak strzałek między niektórymi zmiennymi. I tak, na Rysunku E.1 brak strzałki między F a W_3 oznacza, że – zgodnie z danym rozkładem – F jest niezależne probabilistycznie od W_3 , wzięwszy pod uwagę W_2 . Mówiąc inaczej, fumigacja oddziałuje na populację węgorka tylko w krótkim okresie po jej zastosowaniu, ale pozostaje bez wpływu na jej dalszy rozwój.

Bayesowskie modele przyczynowości pozwalają na dwojakiemu rodzaju wnioskowania. Pierwsze dotyczą sytuacji nieeksperymentalnych, gdy dokonuje się przewidywania w oparciu o związki przyczynowe, zawarte w modelu. Takie wnioskowanie w zasadzie niczym nie odbiega od wnioskowania w sieciach bayesowskich, gdzie zasadniczym twierdzeniem, pozwalającym uaktualniać informacje na podstawie danych doświadczenia, jest twierdzenie Bayesa.²

Drugie dotyczą sytuacji, gdy interwencja na zmienną X modelu wprowadza modyfikację zależności przyczynowych. Przeprowadzenie wnioskowania wymaga wykorzystania tych samych reguł, które stosuje się do przewidywania efektów przyczynowych, ale stosuje się je w zmodyfikowanym grafie, uwzględniającym dokonane zmiany w zależnościach przyczynowych. W graficznej wersji modelu bayesowskiego tę modyfikację w ogólnej postaci można reprezentować poprzez dodanie dodatkowej zmiennej reprezentującej interwencję, a następnie przez stosowanie kryterium d-separacji (por. Punkt 4.4).

Przyjmując, że przed interwencją i po niej pozostałe związki przyczynowe nie uległy zmianie oraz że w obu przypadkach mamy te same czynniki zakłócające, można przewidzieć efekty określonego działania (Pearl 2000, 219). I tak, w podanym przykładzie, aby określić całkowity wpływ fumigacji na wydajność upraw owsa przy ustaleniu fumigacji na poziomie $F=f$ przez interwencję, pomija się zależność funkcyjną F od W_0 , a w pozostałych zależnościach zastępuje się F przez $F=f$. Rezultatem tego kontrfaktycznego wnioskowania jest Równanie (E.1), które określa wpływ fumigacji na uprawy owsa, w sposób zależny tylko od danych doświadczenia.

² Por. (Kawalec 2003b). Oryginał tekstu T. Bayesa zamieszczony jest wraz z bogatą dyskusją w (Swinburne 2002).

F: Modele przyczynowe BN jako uogólnienie modeli SEM

Opracowana przez Wrighta w 1919 roku metoda analizy ścieżek pozwalała ustalać zależności przyczynowe na podstawie pomiarów korelacji, o ile diagram ścieżek poprawnie je reprezentował. Metoda Wrighta polegała na wypisaniu równań – jednego dla każdej pary zmiennych (np. X_i, X_j) i zrównaniu (standaryzacji) współczynnika korelacji ρ_{ij} z sumą iloczynów współczynników ścieżek oraz korelacji resztkowych dla różnych ścieżek łączących X_i z X_j . Można następnie próbować rozwiązać te równania dla współczynników ścieżek w kategoriach zaobserwowanych korelacji. Gdy otrzymane równania dają jednoznaczne rozwiązanie dla pewnego współczynnika ścieżek ρ_{mn} , który jest niezależny od (niezaobserwowanych) korelacji resztkowych, to można o nim powiedzieć, że jest identyfikowalny. Jeśli każdy zbiór współczynników korelacji ρ_{ij} jest zgodny z pewnym wyborem współczynników ścieżek, to o takim modelu mówi się, że jest nietestowalny¹, ponieważ można go doskonale dopasować do dowolnych danych.

Te same problemy testowalności i identyfikowalności, które Wright rozwiązuje częściowo graficznie, a częściowo algebraicznie, można (Pearl 2000, 140) za pomocą zorientowanych grafów niecyklicznych zinterpretowanych przyczynowo rozważać czysto graficznie, nawet przed dostarczeniem danych. Ponadto można odnosić te analizy do modeli innych niż liniowe, a więc nieliniowych czy nieparametrycznych.

Zestawiając notację sieci bayesowskich ze stosowaną w analizie ścieżek, zauważmy, że każdy graf reprezentuje zbiór równań strukturalnych o postaci: $x_i = f_i(pa_i, \epsilon_i)$, dla $i = 1, \dots, n$.

Zmienna pa_i oznaczająca na grafie rodziców x_i reprezentuje w tym równaniu wartości zbioru zmiennych, o których zakłada się, że są bezpośrednimi przyczynami X_i , natomiast ϵ_i reprezentuje błędy wynikłe z pominięcia innych czynników przyczynowych oddziałujących na X_i .

Równanie o tej postaci jest nieliniową i nieparametryczną generalizacją standardowych równań liniowych

$$x_i = \sum_{k \neq i} \alpha_{ik} x_k + \epsilon_i, \quad i = 1, \dots, n. \quad (\text{F.1})$$

¹ Inne określenia to niefalsyfikowalny i przepelniony.

W poprzednim równaniu zbiór zmiennych pa_i odpowiada tym zmiennym z prawej strony powyższego równania, które mają niezerowe współczynniki. Zbiór równań nieparametrycznych, jak podane w powyższym przykładzie, nazywany jest modelem przyczynowym, jeśli każde z tych równań reprezentuje proces przyczynowy, który prowadzi do określenia wartości zmiennej X_i , a nie tylko jej prawdopodobieństwa. Graf otrzymany przez wykreślenie strzałek od każdego elementu pa_i do X_i nazywa się diagramem przyczynowym. Oprócz pełnych jednokierunkowych strzałek diagram przyczynowy powinien zawierać także dwukierunkowe łuki pomiędzy każdą parą zmiennych, których błędy losowe są zależne probabilistycznie.

Jak akcentuje Pearl,

ważne jest, by podkreślić, że diagramy przyczynowe (podobnie jak tradycyjne diagramy ścieżek) powinny się odróżniać od bogatej różnorodności modeli graficznych, jakie spotykamy w literaturze statystycznej, a których konstrukcja i interpretacja opiera się całkowicie na rozkładzie łącznym (Kiiveri i in. 1984; Whittaker 1990; Cox i Wermuth 1996; Lauritzen 1996; Andersson i in. 1999). Brakujące krawędzie w tych modelach statystycznych są reprezentacją warunkowych niezależności, podczas gdy brakujące krawędzie w diagramach przyczynowych – niewystępowania zależności przyczynowych, które mogą, choć nie muszą, implikować warunkowe niezależności w rozkładzie. (2000, 141)

Jeśli zmienne reprezentujące błędy losowe mają wielowymiarowy rozkład normalny, jak powszechnie zakłada się przy modelowaniu równań strukturalnych, to zmienne X_i podane w powyższym równaniu liniowym będą miały też taki rozkład, a w konsekwencji będą w pełni scharakteryzowane przez współczynniki korelacji ρ_{ij} . Bardzo użyteczną własnością wielowymiarowych rozkładów normalnych jest to, że wariancja warunkowa $\sigma_{X|z}^2$, warunkowa kowariancja $\sigma_{XY|z}$ oraz warunkowy współczynnik korelacji $\rho_{XY|z}$ są niezależne od wartości z . Na ich określenie używa się terminu „cząstkowa”: cząstkowa wariancja, cząstkowa kowariancja oraz współczynnik korelacji cząstkowej, i oznacza się je odpowiednio przez: σ_{XZ} , σ_{XYZ} oraz ρ_{XYZ} , gdzie X i Y są pojedynczymi zmiennymi, natomiast Z reprezentuje zbiór zmiennych. Ponadto współczynnik korelacji cząstkowej ρ_{XYZ} jest równy zero wtedy i tylko wtedy, gdy warunek ($X \perp\!\!\!\perp Y | Z$) jest spełniony w rozkładzie.

Współczynnik regresji dany jest więc wzorem

$$r_{XYZ} = \rho_{YXZ} \frac{\sigma_{YZ}}{\sigma_{XZ}}. \quad (\text{F.2})$$

Jest zatem równy współczynnikowi X w liniowej regresji Y na X oraz Z . Innymi słowy, współczynnik x w równaniu regresji

$$y = ax + b_1z_1 + \dots + b_kz_k \quad (\text{F.3})$$

jest określony przez wzór

$$a = r_{YXZ_1Z_2\dots Z_k}. \quad (\text{F.4})$$

Zgodnie z twierdzeniem H. Crámera (1946), współczynniki te można estymować za pomocą metody najmniejszych kwadratów.

Z założenia Markowa wynika, że cząstkowa korelacja ρ_{YXZ} znika w każdym modelu Markowa mającym tę samą strukturę co graf wyjściowy, gdy węzły odpowiadające zmiennym w Z d-separują węzeł X od węzła Y w G , niezależnie od parametrów modelu. Ponadto żadna inna korelacja cząstkowa nie znika dla wszystkich parametrów modelu. W ten sposób otrzymuje się graficzny test na sprawdzenie zależności zachodzących między zmiennymi w modelu.

W systemach ograniczonych, tzn. takich, gdzie każda ukryta zmienna ma ograniczony wpływ na co najwyżej dwie zmienne zaobserwowane (por. Spirtes i in. 1996), kryterium d-separacji implikuje, że te i tylko te korelacje cząstkowe, które stwierdza się przez sprawdzenie d-separacji, mają gwarancję zniknięcia niezależnie od parametrów α_{ik} modelu oraz niezależnie od wariancji błędów losowych. Dzięki temu, jak podkreśla Pearl (2000, 142), mamy prostą metodę testowania modeli, a mianowicie testujemy bezpośrednio każdą zerową korelację cząstkową, którą ten model implikuje. W ten sposób nie trzeba szukać maksymalnego estymatora wiarygodności MLE dla parametrów modelu ani dopasować takiego estymatora do danych.

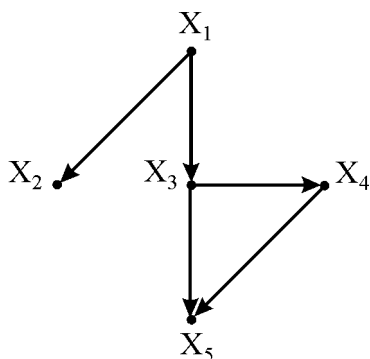
Liczba takich korelacji może być duża, więc istotne jest twierdzenie, że można je wyprowadzić ze stosunkowo niewielkiej liczby cząstkowych korelacji, które konstituują bazę wszystkich pozostałych (Pearl i Verma 1987).

Jako narzucający się wybór bazy dla wyznaczenia zbioru zerowych korelacji cząstkowych, wynikających z danego grafu D , jest zbiór równości $B = \{\rho_{ij-pa_i} = 0 \mid i > j\}$, gdzie i reprezentuje wszystkie węzły w D , a j reprezentuje wszystkie poprzedniki i w dowolnej kolejności, która zgadza się z

kierunkiem strzałek w D .² Testowanie tych równości jest więc warunkiem wystarczającym sprawdzenia wszystkich twierdzeń statystycznych danego liniowego modelu z własnością Markowa. Gdy zbiór rodziców PA_i jest duży, możliwe jest wybranie bardziej ekonomicznej bazy (Pearl 2000, 143): Niech (i, j) oznacza parę niesąsiadujących węzłów w grafie D oraz niech Z_{ij} oznacza dowolny zbiór węzłów, które są bliżej i niż j , oraz takich, że Z_{ij} d-separuje i od j . Zbiór zerowych korelacji cząstkowych $B = \{\rho_{ij \cdot pa_i} = 0 \mid i > j\}$, składający się z jednego elementu dla każdej niesąsiedniej pary, konstituuje bazę dla zbioru wszystkich zerowych korelacji cząstkowych, które wynikają z grafu D .

Powyższe stwierdzenie wyraża zależność między zbiorem zerowych korelacji cząstkowych odpowiadających d-separacji między niesąsiednimi węzłami diagramu a informacją statystyczną zawartą w liniowym modelu z własnością Markowa.

Pearla pojęcie bazy zilustruję zależnościami w diagramie przedstawionym na Rysunku F.1.



Rysunek F.1. (Opracowano na podstawie: Pearl 2000).

Dla tego diagramu dopuszczalna w zasadzie jest jedna z dwóch baz (każdy z elementów tych baz odpowiada brakującej strzałce na diagramie):

$$B_1 = \{\rho_{32 \cdot 1} = 0, \rho_{41 \cdot 3} = 0, \rho_{42 \cdot 3} = 0, \rho_{51 \cdot 43} = 0, \rho_{52 \cdot 43} = 0\}, \quad (F.5)$$

$$B_2 = \{\rho_{32 \cdot 1} = 0, \rho_{41 \cdot 3} = 0, \rho_{42 \cdot 1} = 0, \rho_{51 \cdot 43} = 0, \rho_{52 \cdot 1} = 0\}. \quad (F.6)$$

² Ten zbiór równości reprezentuje cechę „przesiewu względem rodziców” w modelu z własnością Markowa (Pearl 2000, 143).

Baza B_2 jednak wykorzystuje znacznie mniejszy zbiór zależności przy d-separacji i w związku z tym prowadzi do testów, które wymagają sprawdzenia mniejszej liczby regresorów. Minimalna ich liczba równa jest ilości brakujących krawędzi na diagramie, a więc im bardziej przerzedzony jest dany diagram, tym bardziej ogranicza macierz kowariancji i tym więcej trzeba przeprowadzić testów, by zweryfikować te ograniczenia (por. też Pearl 2000, 143-44).

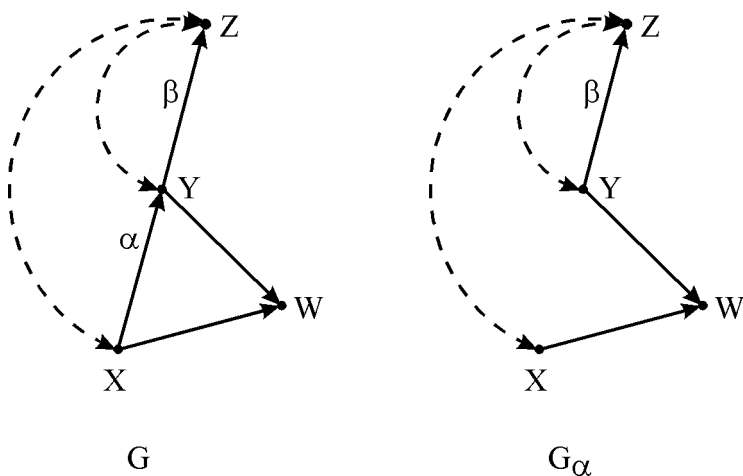
Rozważmy współczynnik regresji dla ścieżki $X \rightarrow Y$. Niech α oznacza współczynnik ścieżki dla tego przypadku. Współczynnik regresji

$$r_{XY} = \rho_{YX} \frac{\sigma_Y}{\sigma_X} \quad (\text{F.7})$$

można rozłożyć na sumę odpowiedniego współczynnika ścieżki oraz cząstkową korelację między X a Y , która powstanie po ustawieniu wartości α na zero (odpowiednik grafu, w którym nie ma krawędzi od X do Y):

$$r_{XY} = \alpha + I_{YX}. \quad (\text{F.8})$$

Oczywiście, jeśli w grafie, w którym usunięto krawędź $X \rightarrow Y$, korelacja między X a Y będzie zerowa, wówczas $I_{YX} = 0$ oraz $r_{XY} = \alpha$. Zachodzenie takiej zależności można sprawdzić graficznie poprzez określenie, czy zmienna X jest d-separowana od zmiennej Y . W przypadku, gdy $r_{XY} = \alpha$, α jest zidentyfikowane, a ilustrację takiego testu identyfikowalności dla prostego przypadku podaje Rysunek F.2.



Rysunek F.2. (Opracowano na podstawie: Pearl 2000).

Na diagramie G_α przedstawionym na Rysunku F.2 wszystkie ścieżki pomiędzy X a Y są zablokowane przez zbieżne strzałki, a więc można obliczyć α na podstawie r_{XY} .

Pearl uogólnia tę zasadniczą ideę na przypadki, gdzie I_{XY} jest różne od zera, chociaż może być sprowadzone do tej wartości dla pewnego zbioru zmiennych $Z = \{Z_1, Z_2, \dots, Z_k\}$, które znajdują się na różnych d-połączonych ścieżkach między X a Y .

Korzystając z powyższej ilustracji, rozważmy współczynnik regresji cząstkowej

$$r_{XYZ} = \rho_{YXZ} \frac{\sigma_{YZ}}{\sigma_{XZ}}, \quad (\text{F.9})$$

który reprezentuje resztkową korelację pomiędzy Y a X po dokonaniu analizy korelacji cząstkowych dla Z . Jeśli w Z nie ma żadnych potomków Y , to ten współczynnik dany jest wzorem (por. Pearl 2000, 150 p. 12):

$$r_{XYZ} = \alpha + I_{YXZ}, \quad (\text{F.10})$$

gdzie I_{YXZ} reprezentuje cząstkową korelację pomiędzy X a Y , powstałą przez ustawienie α na zero, a więc grafu, który będzie się różnił od wyjścio-

wego tak, jak G_α od G na Rysunku F.2, czyli pominięta zostanie w nim krawędź $X \rightarrow Y$.

W przypadku, gdy zmienne ze zbioru Z d-separują X od Y w grafie, gdzie pominięto krawędź $X \rightarrow Y$, $I_{YX:Z}$ będzie równe zero, a tym samym można wywnioskować, że

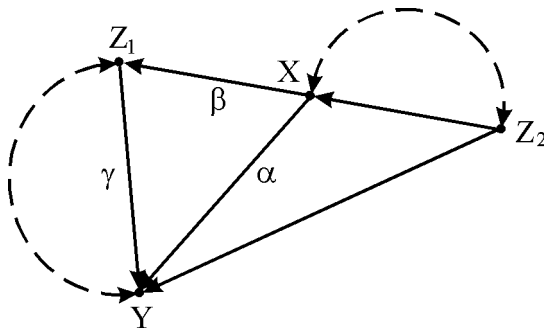
$$r_{XYZ} = \alpha. \tag{F.11}$$

A ponieważ r_{XYZ} jest określone w tym przypadku przez α , czyli współczynnik x w równaniu regresji Y na X i Z , więc α można estymować, wykorzystując równanie regresji

$$y = \alpha x + \beta_1 z_1 + \dots + \beta_k z_k + \varepsilon. \tag{F.12}$$

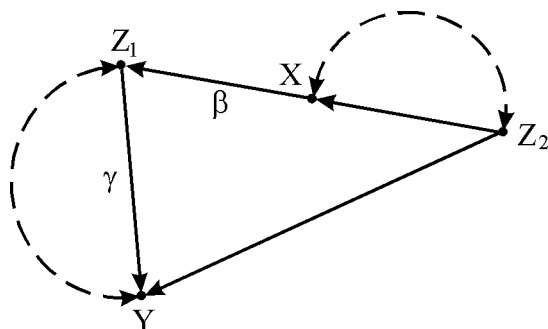
Uogólniając, można stwierdzić, że współczynnik korelacji r_{XYZ} jest zgodnym estymatorem współczynnika ścieżki, jeśli w diagramie, który powstaje z wyjściowego po usunięciu krawędzi $X \rightarrow Y$, istnieje taki zbiór zmiennych Z , że Z nie zawiera potomków Y oraz Z d-separuje X od Y w tym grafie (Pearl 2000, 150-51).

Nie wszystkie jednak przypadki, gdy pytamy nie o bezpośredni efekt przyczynowy X na Y , a o całkowity, mogą być rozpatrywane w ten właśnie sposób. Weźmy pod uwagę graf, przedstawiony na Rysunku F.3.



Rysunek F.3. (Opracowano na podstawie: Pearl 2000).

Jeśli z tego grafu usunąć krawędź $X \rightarrow Y$, to nie będzie takiego zbioru węzłów Z , który d-separowałby wszystkie ścieżki od X do Y .



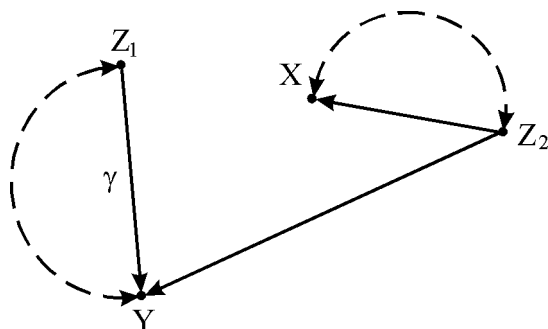
Rysunek F.4. (Opracowano na podstawie: Pearl 2000).

Jeśli by o grafie przedstawionym na Rysunku F.4 założyć, że zbiór Z zawiera Z_1 , to wówczas ścieżka $X \rightarrow Z_1 \leftrightarrow Y$ zostanie odblokowana przez strzałki zbieżne w Z_1 . Jeśli natomiast Z nie będzie zawierać Z_1 , wówczas odblokowana będzie ścieżka $X \rightarrow Z_1 \rightarrow Y$. Prowadzi więc to do wniosku, że nie można zidentyfikować współczynnika ścieżki α w sposób opisany powyżej.

Pearl (2000, 151) sugeruje więc inny sposób, poprzez tzw. kryterium „tylnych drzwi” (por. Równanie 4.40). Całkowity efekt przyczynowy X na Y wyrażony jest przez $\alpha + \beta\gamma$, a w celu zidentyfikowania tej sumy dzięki r_{YZ} ten współczynnik nie może „uczestniczyć” w innych ścieżkach niż te, które prowadzą od X do Y . Jednak w grafie na Rysunku F.4 są dwie takie ścieżki: $X \leftarrow Z_2 \rightarrow Y$ oraz $X \leftrightarrow Z_2 \rightarrow Y$. Takie ścieżki Pearl określa mianem ścieżek „tylnych drzwi” lub konfundujących. Obie te ścieżki są jednak blokowane przez Z_2 , a więc wzięcie poprawki na tę zmienną pozwoli na zidentyfikowanie $\alpha + \beta\gamma$. Otrzymamy więc

$$\alpha + \beta\gamma = r_{XY.Z_2}. \quad (\text{F.13})$$

Dla analizowanego przykładu utworzenie podgrafu $G_{\underline{X}}$ (Rysunek F.5) ilustruje, że spełnione jest kryterium „tylnych drzwi”.



Rysunek F.5. (Opracowano na podstawie: Pearl 2000).

Po uwzględnieniu zmiennej Z_2 na powyższym grafie nie ma konfundującej ścieżki, która łączyłaby zmienne X oraz Y , a zatem całkowity efekt przyczynowy w tym przypadku dany jest wzorem $r_{XY;Z_2}$. Można więc na podstawie analizy tego przykładu wysunąć wniosek, że całkowity efekt przyczynowy został określony na podstawie grafu bez konieczności identyfikowania wszystkich składowych tego grafu. Pearl podkreśla tę własność wnioskowania opartego na graficznej metodzie reprezentacji jako jeden z ważnych elementów uogólnienia dotychczasowych modeli równań strukturalnych (2000, 152).³

Taka interpretacja równań strukturalnych pozwala Pearlowi na redefinicję pojęcia egzogeniczności. Oznaczmy przez λ zbiór parametrów pointerwencyjnego prawdopodobieństwa $P(y \parallel x)$, a przez T wiedzę teoretyczną w odniesieniu do danego modelu, w którym występują zmienne X oraz Y . X jest egzogeniczna względem układu (Y, λ, T) , jeśli λ jest identyfikowalne na podstawie rozkładu warunkowego $P(y \mid x)$, a więc gdy dla dowolnych dwóch modeli M_1 oraz M_2 zgodnych z T spełniony jest warunek:

$$P_{M_1}(y \mid x) = P_{M_2}(y \mid x) \rightarrow \lambda[P_{M_1}(y \mid x)] = \lambda[P_{M_2}(y \mid x)].^4 \quad (F.14)$$

³ Szczegółową dyskusję przypadków, kiedy nie można zidentyfikować całkowitego efektu przyczynowego bez zidentyfikowania wszystkich poszczególnych komponentów grafu Pearl przeprowadza w (2000, 152-154).

⁴ (Pearl 2000, 166). Uogólnioną i semantycznie zorientowaną postać tej definicji podaje i konfrontuje z innymi propozycjami Pearl (2000, 168-70).

Szczególnymi przypadkami tej definicji są sytuacje, gdy można określić pełną charakterystykę pointerwencyjnych prawdopodobieństw i wówczas warunek F.14 przyjmie uproszczoną postać:

$$P_{M_1}(y | x) = P_{M_2}(y | x) \rightarrow P_{M_1}(y || x) = P_{M_2}(y || x). \quad (F.15)$$

Ponadto gdy teoria dopuszcza jedynie modele o tej samej strukturze grafu, wówczas mamy kolejne uproszczenie tego warunku:

$$P(y || x) = P(y | x). \quad (F.16)$$

Oznacza to, że jeśli X jest zmienną egzogeniczną w sensie Pearla, to efekty przyczynowe interwencji na tę zmienną można przewidzieć „bezpośrednio z pasywnych obserwacji, nawet bez uwzględniania czynników konfundujących” (Pearl 2000, 167). To stwierdzenie jest niezwykle ważne z kilku względów. Pearla definicja egzogeniczności⁵ jest wyraźnie zrelatywizowana do teorii czy jakościowej wiedzy ogólnej, która określa pewne zależności przyczynowe i która oczywiście podlega, jak zwraca na to uwagę Freedman (1997a), niezależnemu sprawdzeniu.

Ta definicja w klarowny sposób określa także jeden z istotnych aspektów interwencji, jakie ma ona w badaniach nieeksperymentalnych. Otóż, jak wyżej zostało podkreślone, w przypadku zmiennej egzogenicznej efekt przyczynowy interwencji na tę zmienną można przewidzieć wyłącznie na podstawie danych uzyskanych w badaniach nieeksperymentalnych. Oczywiście pamiętamy o zastrzeżeniu poczynionym wyżej, a które notorycznie w swoich pracach podkreśla Freedman – stwierdzenie egzogeniczności nie dokonuje się, jak postulują m.in. D. Hausman i Woodward (2004), *wyłącznie* względem modelu, lecz także względem leżących u jego podstaw założeń, które wymagają odrębnej weryfikacji.

⁵ Inne znane w literaturze definicje egzogeniczności podają m.in. Engle i in. (1983), Leamer (1985), Maddala (1992), Darnell (1994), Hendry (1995), Freedman (1997a), D. Hausman i Woodward (2004).

G: Przykłady zastosowań modeli BN do ustalania przyczyn

Niektórzy badacze (np. Lieberson 1997; Freedman 1997a; Robins 1999) są bardzo ostrożni w przyjmowaniu modelowania przyczynowego za pomocą sieci bayesowskich i posługiwaniu się nim w badaniach. Ten sceptycyzm motywują również dotychczasowym brakiem zastosowań tego sposobu modelowania w rozstrzyganiu zasadniczych hipotez naukowych. Wydaje się, że jest jeszcze za wcześnie, by ostatecznie móc zweryfikować ten sposób modelowania przyczynowości, powołując się na aplikacje lub na ich ewentualny brak. Realizowany obecnie w London School of Economics rozbudowany projekt opracowania zastosowań manipulacyjnego podejścia do przyczynowości w różnych naukach (w tym przyrodniczych i społecznych), inspirowany głównie pracami Pearla, nie zakończył się jeszcze. Ta teoria jest także daleka od powszechności w wykładzie podręcznikowym.

Bez ambicji ostatecznego rozstrzygnięcia powyższego dylematu przestudiuję poniżej kilka znanych z literatury metodologicznej aplikacji sieci bayesowskich do modelowania przyczynowego.¹ Niektóre z nich doczekały się już obfitych komentarzy.

Jedną z dyscyplin, gdzie zdaniem biologa ewolucyjnego B. Shipleya (1999) sieci bayesowskie znajdują zastosowanie w analizie zależności przyczynowych, jest biologia. Jest to do pewnego stopnia zaskakujące, gdyż poprzednik tej metody, a mianowicie modele równań strukturalnych nie znalazły miejsca w badaniach w tej dyscyplinie. Shipley podaje frapujące wyjaśnienie tego faktu:

Dyscypliny takie, jak socjologia czy psychologia, które także [obok biologii organizmów] nie posiadają ścisłej bazy teoretycznej, ale które wprowadziły modelowanie równań strukturalnych, zajmują się przynajmniej organizmem, z którym wszyscy mamy bliską znajomość [ang. know intimately]; intuicyjne odwołanie się do „przyczyn” nie wydaje się więc tak szokujące, gdy badamy samych siebie. Biolog badający koniki polne czy drzewa klonowe nie może odwoływać się do takich zastępczych doświadczeń. (1999, 442)

Za główny problem, którego rozwiązanie mogą ułatwić metody modelowania przyczynowości za pomocą sieci bayesowskich, Shipley uznaje konstruowanie modeli, które następnie można poddać testowaniu empirycz-

¹ Ilustracją zastosowania nieprzyczynowego modelu bayesowskiego w badaniach psychologicznych jest m.in. praca A. Bieli i in. (1974).

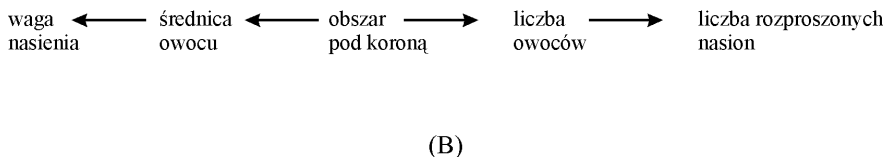
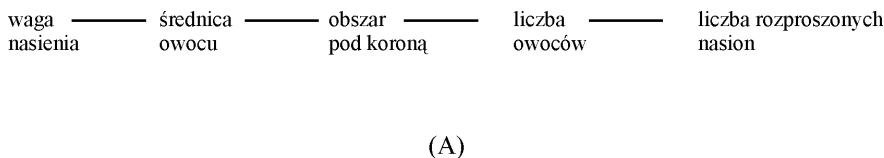
nemu. W odróżnieniu od metod statystycznych bowiem pozwalają one na przeszukiwanie przestrzeni możliwych modeli, a nie tylko ich testowanie.

Z jednej strony, cechy organizmów często współwystępują, np. ostrym zębom często towarzyszą pazury, silne mięśnie nóg i wyostrome zmysły, co można uzasadniać ewolucyjnie jako efekt maksymalizacji dostosowania w procesie doboru naturalnego. Ten sam proces jednak wymaga niekiedy zrównoważenia przeciwstawnych cech.

Pierwszy z problemów, który podejmuje Shipley, dotyczy względnej liczby nasion rozpraszanych, zwłaszcza przez ptaki, poszczególnych egzemplarzy drzewa o nazwie St. Lucie's Cherry (*Prunus mahaleb*). Pięć uwzględnionych w badaniach zmiennych to: obszar pod koroną drzewa (miara całkowitej biomasy fotosyntetycznej), całkowita liczba dojrzałych owoców na drzewo, średnia średnica owocu, średnia waga nasienia oraz liczba nasion rozproszonych z drzewa przez ptaki. Wykorzystano dane o 60 drzewach; obszar przekształcono w pierwiastek kwadratowy, a pozostałe zmienne w logarytmy naturalne dla zachowania normalności ich rozkładu.

Celem tego badania było ustalenie modelu najlepiej dostosowanego do kowariancji tych pięciu cech oraz najlepiej przewidyującego ilościowe wartości każdego z atrybutów. Wyniki otrzymane dzięki zastosowaniu algorytmu PC programu TETRAD przedstawia Rysunek G.1.

Graf na Rysunku G.1 w panelu (a) przedstawia topologię zależności między tymi atrybutami, z którą zgodnych jest jedenaście różnych modeli ze zorientowanymi krawędziami. W dalszych rozważaniach można posiłkować się podstawową wiedzą w zakresie biologii, przynajmniej przy określeniu kierunku niektórych z tych krawędzi. I tak, liczba nasion rozproszonych przez ptaki w końcowym okresie wzrostu musi – ze względu na informacje o przebiegu czasowym tego procesu – zależeć od liczby owoców na danym drzewie, a nie na odwrót. Ta informacja redukuje liczbę możliwych grafów do siedmiu.



Rysunek G.1. (Opracowano na podstawie: Shipley 1999).

W podobny sposób można wywnioskować, że całkowita liczba owoców wyprodukowanych przez dane drzewo jest determinowana przez całkowite rozmiary korony drzewa, a więc przez całkowite zasoby, jakie ma do dyspozycji przy produkcji owoców. Wiadomo bowiem, że produkty fotosyntezy oraz substancje odżywcze są kierowane z liści lub organów je magazynujących do rozwijającego się owocu czy nasienia. Ponadto brak zasobów spowodowałby zaprzestanie rozwoju nasion, które zaczęły się formować. Tę informację można wykorzystać do zorientowania kolejnej krawędzi i zredukowania liczby możliwych modeli do czterech.

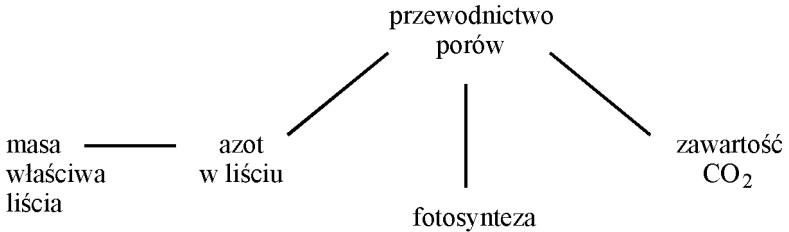
Analogiczne rozumowanie pozwala zorientować krawędź łączącą obszar pod koroną i średnią średnicę owocu. Pozostają więc tylko dwa grafy.

Za najbardziej problematyczną uznaje Shipley krawędź między średnicą owocu a wagą nasienia. Powołując się na przeprowadzone obliczenia, stwierdza, że skierowanie jej do wagi nasienia daje model nieco lepiej dopasowany do danych. Nie jest to jednak, zdaniem Shipleya, wystarczający powód, by wybrać jeden z nich. Podsumowując, deklaruje on, że ten przykład

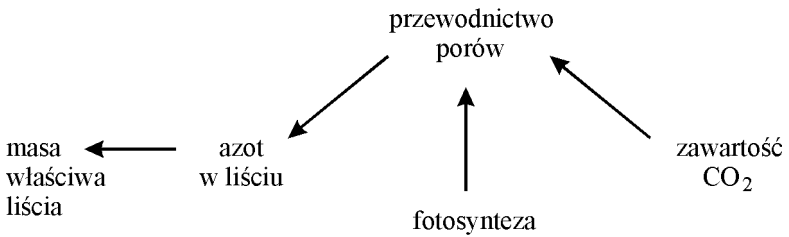
Aplikacja [...] algorytmu SGS doprowadziła do utworzenia modelu, który zgadza się zarówno z dostępnymi danymi, jak i ze znanymi wcześniej rezultatami empirycznymi. Ponieważ ma obie te cechy, więc jest, moim zdaniem, wart testowania w stosunku do niezależnych danych. (1999, 447)

Drugi przykład dotyczy fotosyntezy, czyli przetwarzania przez rośliny dwutlenku węgla pobieranego z atmosfery w proste węglowodany. W tym przypadku rośliny muszą powiększyć rozmiar niewielkich otworów na powierzchni liści, aby wpuścić dwutlenek węgla do przestrzeni międzykomórkowej.

kowej. Wraz jednak z poszerzeniem tych otworów woda z wilgotnych tkanek przemieszcza się do zwykle bardziej suchej atmosfery. Fotosynteza łączy się zatem z problemem maksymalizacji pobieranego dwutlenku węgla i jednocześnie minimalizacji utraty wody. W „rozwiązywaniu” tego problemu w danym kontekście środowiskowym uczestniczą różne cechy liści. Zadanie teoretyczne polega więc na skonstruowaniu matematycznego modelu, który opisywałby, jak te cechy wzajemnie na siebie oddziałują w rozwiązywaniu tego problemu.



(A)



(B)

Rysunek G.2. (Opracowano na podstawie: Shipley 1999).

Do cech, które są zaliczane jako relewantne, należą: masa właściwa liścia (powierzchnia liścia podzielona przez wagę liścia po wysuszeniu), całkowita koncentracja azotu w liściach, przewodnictwo wodne porów, tempo fotosyntezy oraz koncentracja dwutlenku węgla w liściu. Dane, którymi posłużył się Shipley, zawierają trzydzieści pięć obserwacji, z których każda składa się ze średniej dla każdej z pięciu zmiennych dla różnych gatunków roślin –

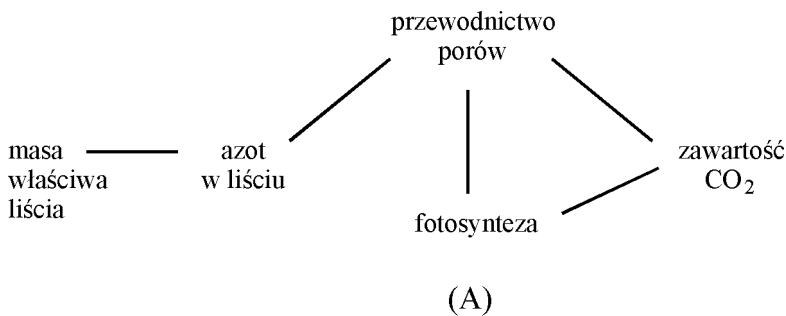
średnie oparte są na około dwunastu roślinach dla jednego gatunku. Każda wartość normalna jest znormalizowana, tzn. przekształcona do postaci logarytmu naturalnego.

Panel (A) na Rysunku G.2 przedstawia topologię grafu, którą otrzymano dla najczęściej powtarzających się krawędzi. Z dwóch otrzymanych acyklicznych grafów skierowanych panel (B) przedstawia ten lepiej dopasowany do danych. Shipley podkreśla jednak, że żaden z dwóch alternatywnych modeli nie jest zgodny z obowiązującą teorią i oba zostaną odrzucone przez biologów.

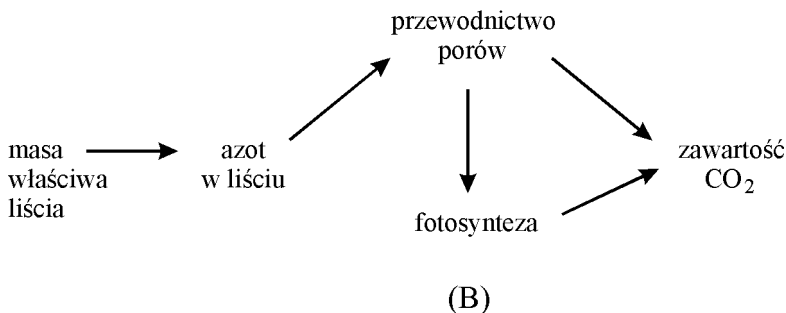
Shipley tłumaczy to niespełnieniem założenia wierności. Zawartość dwutlenku węgla i fotosyntezę łączą alternatywne ścieżki o przeciwnych kierunkach, których efekt w zasadzie znosi się wzajemnie. Przy narzuceniu algorytmowi PC ścieżki od przewodnictwa porów liścia przez fotosyntezę do zawartości dwutlenku węgla – ścieżek, których istnienia wymaga elementarna fizyka zjawiska dyfuzji oraz biochemia fotosyntezy – Shipley otrzymał rezultat przedstawiony na Rysunku G.3.

Z topologią grafu przedstawionego na Rysunku G.3 w panelu (A) powyżej zgodnych jest pięć modeli, które są zgodne z danymi. Uwzględnienie elementarnych praw fizyki, fizjologii i biochemii fotosyntezy prowadzi do grafu przedstawionego na panelu (B). Taki model został niezależnie opracowany przez M. Lechowicza i Shipleya i zapewnia dobre dopasowanie do danych ($X^2 = 4,719$, $p = 0,45$).

Shipley uznaje więc za zaletę manipulacyjnego podejścia do przyczynowości to, że pozwala na przeszukiwanie ogromnej liczby możliwych modeli (np. dla analizowanych tu obu przykładów liczba wszystkich kombinacji to 59 000) i określenie, które z nich najlepiej pasują do danych i przewidują. Podkreślimy, że wniosek Shipleya nie jest jednak jednoznaczny. Pewne metody, jak np. symulacje Monte Carlo, pokazują, że algorytmy SGS skutecznie przeszukują ogromne przestrzenie hipotez i określają modele bliskie prawdziwemu. Ale, jak kontynuuje Shipley,



(A)



(B)

Rysunek G.3. (Opracowano na podstawie: Shipley 1999).

Niemożliwe jest jednak określenie, czy jest to także prawdą w odniesieniu do rzeczywistych danych, ponieważ – z definicji – „prawdziwy” model jest nieznan. W najlepszym razie możemy spodziewać się, że modele otrzymane za pomocą tych algorytmów przeszukiwania będą potrafiły przewidywać schematy kowariancji w niezależnych danych, w tym otrzymanych po manipulacji, i będą zgodne z zależnościami przyczynowymi, jeśli je poznamy. (1999, 451-52)

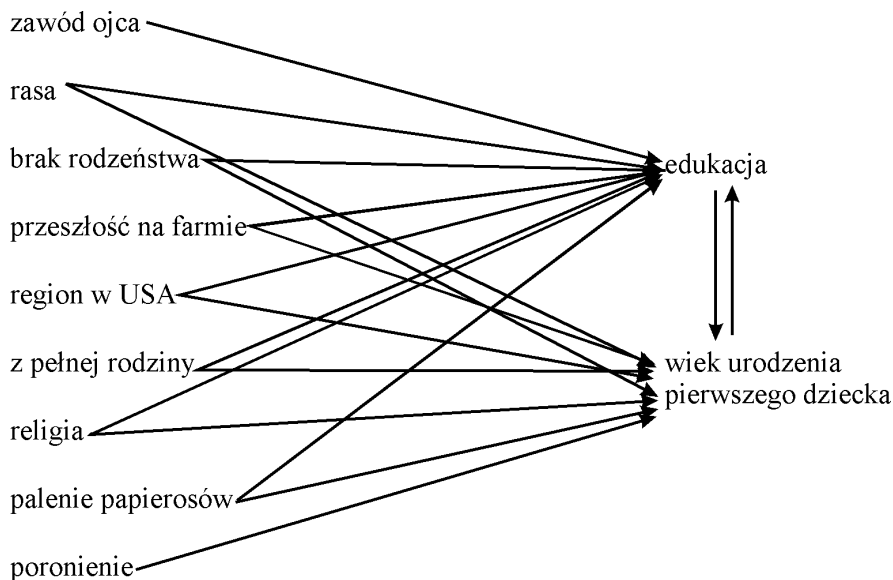
Innego rodzaju przykładem posługują się Spirtes, Glymour i Scheines w demonstrowaniu sprawności programu TETRAD, który sprawdzają m.in. na przykładzie opisanym w literaturze w oparciu o rzeczywiste dane (Spirtes, Glymour i Scheines 2000, 102-103).

Rindfuss i in. (1980) przeprowadzili badania dotyczące zależności wśród zamężnych kobiet między edukacją w czasie małżeństwa (ilość ukończonych lat studiów lub innego rodzaju wykształcenia) a wiekiem kobiety, w którym rodzi ona pierwsze dziecko. Na gruncie teoretycznym uzasadniają model przedstawiony na poniższym rysunku. Wiele spośród zmiennych jest

binarnych: rasa (czarna lub nie), przeszłość na farmie (to, czy dana osoba wychowywała się w dzieciństwie na farmie), region (w którym wychowywała się dana osoba – południe lub inny rejon Stanów Zjednoczonych), pełna rodzina (określane dla wieku 14 lat badanej osoby), religia (katolicka lub inne wyznanie), palenie (przed osiągnięciem 16 lat) i poronienie (przed urodzeniem pierwszego dziecka).

Dane w badaniach przeprowadzonych przez Rindfussa i in. pochodziły od 1766 kobiet w wieku 35-44 lat, które są rezydentami na kontynentalnym obszarze USA. Próbkę ograniczono do zamężnych kobiet, które mają przynajmniej jedno dziecko.

Jak podkreśla Freedman, jednym z kluczowych założeń wyjściowego studium Rindfussa jest brak krawędzi zawód ojca \rightarrow wiek oraz poronienie \rightarrow edukacja, gdyż wymaga tego zastosowana przez nich technika estymacji (dwuetapowa metodą najmniejszych kwadratów). Bez tych lub podobnych założeń system nie byłby identyfikowalny (Freedman 1997a, 124).

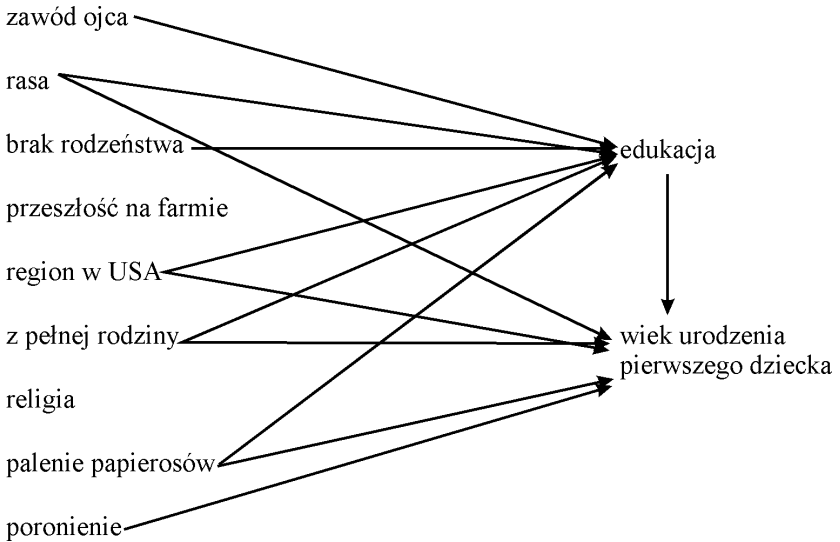


Rysunek G.4. (Opracowano na podstawie: Spirtes, Glymour i Scheines 2000).

Nieoczekiwanym wynikiem badań było stwierdzenie – po estymacji współczynników, że parametr ścieżki wiek \rightarrow edukacja nie jest statystycz-

nie istotny. A więc że te kobiety, które przerywają edukację, zrobiłyby to niezależnie od wieku, kiedy rodzą pierwsze dziecko.

Spirtes, Scheines i Glymour twierdzą, że po wprowadzeniu do programu informacji o tym, że wiek oraz edukacja nie są przyczynami innych zmiennych, używając poziomu istotności 0,5 w algorytmie PC, utworzyli graf przedstawiony na Rysunku G.5.



Rysunek G.5. (Opracowano na podstawie: Spirtes, Glymour i Scheines 2000).

W komentarzu do otrzymanego wyniku Spirtes, Glymour i Scheines podkreślają, że „niektóre czynniki mające wpływ na rzetelność wyników można ocenić na podstawie wiedzy ogólnej” (2000, 351). I tak, w przypadku krawędzi, o której wiadomo, że nie istnieje (gdyż np. przeczy uprzedniości czasowej), można ocenić wynik jako błędny. W analizowanym tu przypadku błąd dotyczył tego, że algorytm PC uruchomiono z założeniem liniowości. A mimo to na otrzymanym w efekcie grafie były krawędzie od zmiennych dyskretnych (np. rasa lub to, czy dana osoba spędziła dzieciństwo na farmie) do zmiennych ciągłych (wiek lub edukacja), co jest w oczywisty sposób niezgodne z przyjętym założeniem (i dlatego można je pominąć jako błędne, jak zrobiono to na Rysunku G.5).

Szczegółowej analizie przykład ten poddaje Freedman (1997a) i Humphreys (Freedman i Humphreys 1996, 1999). Jest ona elementem ostrej krytyki programu automatyzacji wnioskowania, które wykorzystując modele

statystyczne, ma przejść od korelacji do przyczynowości. Freedman i Humphreys deklarują:

identyfikacja zależności przyczynowych wymaga przemyślanej, złożonej i wytrwałej ciężkiej pracy; kluczową rolę odgrywa istotna wiedza naukowa. (1999, 29)

Freedman stawia zasadnicze pytania pod adresem wyjściowej postaci przykładu, jakim posłużyli się Spirtes, Glymour i Scheines. Otóż zastanawia się, czy można, jak przyjął Rindfuss i inni, założyć, że współczynniki są takie same dla wszystkich kobiet – biednych z północnego wschodu po bogatych na przedmieściach zachodu. Zwraca też uwagę, że pominięto istotne zmienne, jak aspiracje, zdolności, zamożność oraz klasa społeczna.

Dla danych z badań Rindfussa i in. Freedman za pomocą programu TETRAD II otrzymał model przyczynowy przedstawiony na Rysunku G.6.² Uznał go za „bardzo dziwny” (1997a, 131). Wynika z niego bowiem, że rasa oraz religia są przyczynami miejsca zamieszkania. Jeśli rozważyć odpowiadające tej zależności równanie:

$$Region = \alpha \cdot Rasa + \beta \cdot Religia + \epsilon, \quad (G.1)$$

to okazuje się, że zmienna *Region* musiałaby przyjąć niedopuszczalne wartości. To jest zmienna binarna, a prawa strona powyższego równania daje wartości ciągle. Jak podkreśla Freedman, tego równania nie można bronić nawet jako przybliżenia (1997a, 132). Pomijając więc kwestie odniesienia do danych spoza tych, które posłużyły do wygenerowania tego modelu, można stwierdzić, że ten model nie spełnia warunku wewnętrznej spójności.

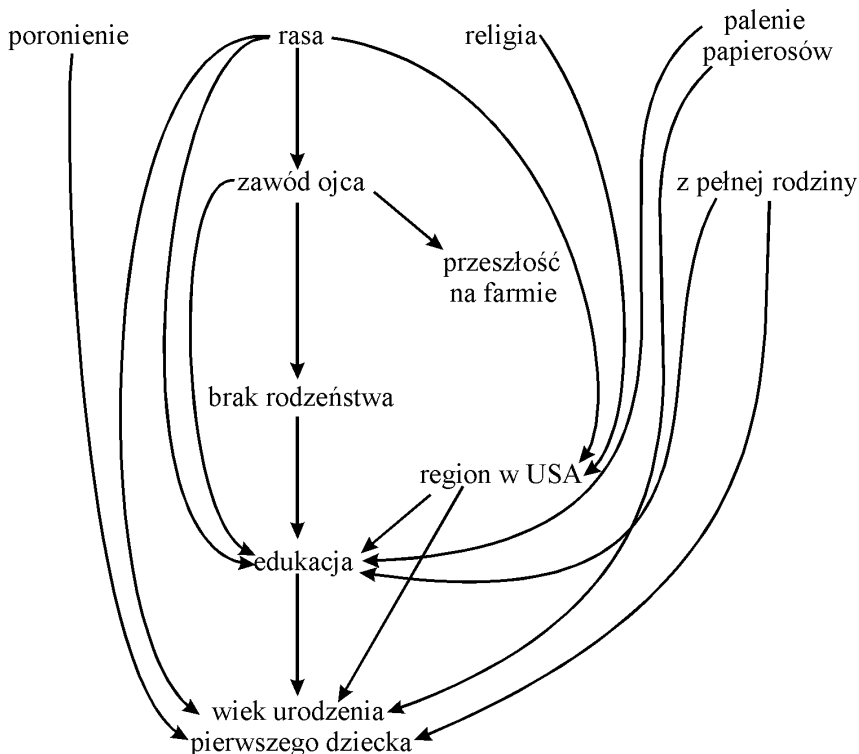
Taki zarzut Freedmana ma jednak odpowiedź, którą już wcześniej zasygnalizowałem, a mianowicie, że w przypadku oczywistego błędu można te zależności pominąć (Spirtes, Glymour i Scheines 2000, 351). Podobnie zarzuty Freedmana odpierają Korb i Wallace (1997, 550-51).

W odpowiedzi Freedman ironizuje:

Te argumenty mają pewien zastanawiający urok. Algorytmy odkrycia przyczyn odnoszą sukces, gdy chroni się je przed popełnieniem błędów. Jeśli algorytmy działają, to nawet *mimo* błędów w założeniach, a jeśli nie – to *ze względu na* błędy w założeniach. [...] TETRAD II wygenerował absurdalne wnioski. Odkrył pewne zależności przyczynowe, które mogą być prawdziwe, ale także odtworzył inne, które są po prostu fałszywe. To nie jest niezawodny

² Szczegółowy opis założeń, z których korzystał Freedman oraz wersji programu, którą wykorzystał, podają Freedman (1997a, 132) oraz Freedman i Humphreys (1999, 42-43).

algorytm, a standardy zastosowane do ich oceny przez Korba i Wallace'a są zanadto elastyczne". (Freedman i Humphreys 1999, 40).



Rysunek G.6. (Opracowano na podstawie: Spirtes, Glymour i Scheines 2000).

Freedman i Humphreys proponują, by ostatecznie tę dyskusję rozstrzygnąć przez podanie „faktycznych przykładów pokazujących sukces algorytmów odkrycia przyczyn” (1999, 40).

Ta krytyka dotyczy tylko fragmentu modelowania przyczynowego, a mianowicie zagadnienia automatyzacji odkrywania zależności przyczynowych na podstawie danych statystycznych. Żaden jednak z podanych przez Freedmana i Humphreysa argumentów nie podważa sensowności stosowania sieci bayesowskich w modelowaniu przyczynowości. Podważą ją jedynie fakt, że sama taka aplikacja najprawdopodobniej nie rozstrzygnie zasadniczych problemów, przy których konieczna będzie także weryfikacja założeń związanych z zastosowaniem modeli sieci bayesowskich badań opartych na rzeczywistych danych.

Bibliografia

- Aalen O. O., 2004, „Discussion of causality”, *Scandinavian Journal of Statistics* 31, s. 193-95.
- Abbott A., 1992, „What do cases do? Some notes on activity in sociological analysis”, w: C. C. Ragin, H. S. Becker (red.), *What is a case?*, Cambridge: Cambridge University Press, s. 53-82.
- Abbott A., 1998, „The causal devolution”, *Sociological Methods and Research* 27, s. 148-91.
- Achen C. H., 1987, „As statisticians see us: comments on Freedman’s paper ‘As others see us...’”, *Journal of Educational Statistics* 12, s. 148-50.
- Adams P., Hurd M., McFadden D., Merrill A., Ribeiro T., 2003a, „Healthy, wealthy, and wise? Tests for direct causal paths between health and socioeconomic status”, *Journal of Econometrics* 112, s. 3-56.
- Adams P., Hurd M., McFadden D., Merrill A., Ribeiro T., 2003b, „Response”, *Journal of Econometrics* 112, s. 129-33.
- Adda J., Chandola T., Marmot M., 2003, „Socio-economic status and health: causality and pathways”, *Journal of Econometrics* 112, s. 57-63.
- Ajdukiewicz K., 1964/1985, „Dowód i wyjaśnienie”, w: K. Ajdukiewicz, 1985, *Język i poznanie*, t. II, Warszawa: PWN, s. 401-408.
- Ajdukiewicz K., 1975, *Logika pragmatyczna*, Warszawa: PWN.
- Amsterdamski S., 1983, *Nauka a porządek świata*, Warszawa: PWN.
- Anderson T. W., 2003, *An introduction to multivariate statistical analysis*, Nowy Jork: Wiley.
- Andersson S. A., Madigan D., Perlman M. D., Richardson T. S., 1999, *Graphical Markov models in multivariate analysis*, w: S. Ghosh (red.), *Multivariate analysis, design of experiments and survey sampling*, Nowy Jork: Dekker, s. 187-229.
- Angrist J. D., Imbens G. W., Rubin D. B., 1996, „Identification of causal effects using instrumental variables: rejoinder”, *Journal of the American Statistical Association* 91, s. 444-72.

- Anscombe G. E. M., 1971/1981, „Causality and determination”, w: G. E. M. Anscombe, 1981, *Metaphysics and philosophy of mind. Collected philosophical papers*, t. 2, Oxford: Blackwell, s. 133-47.
- Appendix to report of the Committee for Scientific Inquiries in relation to the cholera epidemic of 1854*, 1855, Londyn: Her Majesty's Stationery Office.
- Arjas E., 2001, „Causal analysis and statistics: a social sciences perspective”, *European Sociological Review* 17, s. 59-64.
- Arjas E., Parner J., 2004a, „Causal reasoning from longitudinal data”, *Scandinavian Journal of Statistics* 31, s. 171-87.
- Arjas E., Parner J., 2004b, „Reply to discussion”, *Scandinavian Journal of Statistics* 31, s. 198-201.
- Armstrong D. M., 1978, *A theory of universals*, t. 1-2, Cambridge: Cambridge University Press.
- Armstrong D. M., 1983, *What is a law of nature?*, Cambridge: Cambridge University Press.
- Armstrong D. M., 1993, „The identification problem and the inference problem”, *Philosophy and Phenomenological Research* 53, s. 421-22.
- Armstrong D. M., 1997, *A world of states of affairs*, Cambridge: Cambridge University Press.
- Arrow K. J., 1962, „The economic implications of learning by doing”, *The Review of Economic Studies* 29, s. 155-73.
- Arystoteles, *Analityki wtóre*, tłum. K. Leśniak, Warszawa: PWN 1990.
- Ash A., Duan N., Normand S.-L. T., 2001, „HSOR special issue on causal inference: introduction”, *Health Services and Outcomes Research Methodology* 2, s. 165-67.
- Backe A., 1999, „The likelihood principle and the reliability of experiments”, *Philosophy of Science Proceedings* 66, s. S354-61.
- Barnes E., 1992, „Explanatory unification and the problem of asymmetry”, *Philosophy of Science* 59, s. 558-71.
- Basmann R. L., 1988, „Causality tests and observationally equivalent representations of econometric models”, *Journal of Econometrics* 39, s. 7-21.
- Bentler P. M., 1987, „Structural modeling and the scientific method: comments on Freedman's critique”, *Journal of Educational Statistics* 12, s. 151-57.
- Berger R., 1998, „Understanding science: why causes are not enough”, *Philosophy of Science* 65, s. 306-32.
- Berk R. A., 1988, „Causal inference for sociological data”, w: N. J. Smelser (red.), *Handbook of sociology*, Newbury Park: Sage, s. 155-72.
- Bernert C., 1983, „The career of causal analysis in American sociology”, *British Journal of Sociology* 24, s. 230-54.
- Biela A., Bujak F., Walesa C., 1974, „Model Bayesa jako narzędzie pomiaru zachowania (badania eksperymentalne dzieci)”, *Studia Psychologiczne* 13, s. 73-93.
- Bigaj T., 1994, „Przyczynowość w Szkole Lwowsko-Warszawskiej”, *Filozofia Nauki* 6, s. 69-81.

- Blalock H. M., 1960/1977, *Statystyka dla socjologów*, Warszawa: PWN.
- Blalock H. M., 1962, „Four-variable causal models and partial correlations”, *American Journal of Sociology* 68, s. 182-94.
- Blalock H. M. (red.), 1971, *Causal models in the social sciences*, Chicago: Aldine.
- Błaszczyk P., 1999, „Związek przyczynowy w *Sporze o istnienie świata*”, *Kwartalnik Filozoficzny* 27, z. 2, s. 69-118.
- Bobryk J., 1992, *Przyczynowość i intencjonalność*, Warszawa: OW WPUW.
- Bocheński J. M., 1954/1993, *Współczesne metody myślenia*, Poznań: „W drodze”.
- Bogen J., 2004, „Analysing causality: the opposite of counterfactual is factual”, *International Studies in the Philosophy of Science* 18, s. 3-26.
- Box G., Jenkins G., 1976, *Time series analysis, forecasting, and control*, San Francisco: Holden Day.
- Boyd R., 1999, „Kinds, complexity and multiple realization ”, *Philosophical Studies* 95, s. 67-98.
- Burnham K. P., Anderson D. R., 2002, *Model selection and inference*, Nowy Jork: Springer.
- Braithwaite R. B., 1955, *Scientific explanation. A study of the function of theory, probability and law in science*, Cambridge: Cambridge University Press.
- Bromberger S., 1966, „Why-questions”, w: R. G. Colodny (red.), *Mind and cosmos*, Pittsburgh: University of Pittsburgh Press, s. 86-111.
- Bronk A., 1988, *Rozumienie – dzieje – język. Filozoficzna hermeneutyka H.-G. Gadamera*, Lublin: RW KUL.
- Bronk A., 1992, „Wielość nauk i jedność nauki”. w: S. Kamiński, *Nauka i metoda. Pojęcie nauki i klasyfikacja nauk*, oprac. A. Bronk, Lublin: TN KUL, s. 345-70.
- Bronk A., 1998, *Zrozumieć świat współczesny*, Lublin: TN KUL.
- Brzeziński J. (red.), 1987, *Wielozmiennowe modele statystyczne w badaniach psychologicznych*, Warszawa: PWN.
- Brzeziński J. (red.), 1989, *Psychologia matematyczna*, Kraków: W UJ.
- Brzeziński J., 1996, *Metodologia badań psychologicznych*, Warszawa: PWN.
- Brzeziński J., 2000, *Badania eksperymentalne w psychologii i pedagogice*, Warszawa: Scholar.
- Brzeziński J., 2005, „Aktualność klasycznych tekstów z psychometrii dla metodologicznej i etycznej poprawności praktyki diagnostycznej odwołującej się do wyników testów psychologicznych”, tłum. E. Hornowska, M. Zakrzewska, w: J. Brzeziński (red.), *Trafność i rzetelność testów psychologicznych. Wybór tekstów*, Gdańsk: GWP, s. 9-25.
- Bunge M., 1959/1968, *O przyczynowości. Miejsce zasady przyczynowej we współczesnej nauce*, tłum. S. Amsterdamski, Warszawa: PWN.
- Bunge M., 1961, „Kinds and criteria of scientific laws”, *Philosophy of Science* 28, s. 260-81.
- Campbell D. T., Stanley J. C., 1966, *Experimental and quasi-experimental designs for research*, Chicago: Rand McNally.

- Carroll J., 1987, „Ontology and the laws of nature”, *Australasian The Journal of Philosophy* 65, s. 261-76.
- Carroll J., 1994, *Laws of nature*, Cambridge: Cambridge University Press.
- Cartwright N., 1993, „In defence of ‘this worldly’ causality: comments on van Fraassen's *Laws and symmetry*”, *Philosophy and Phenomenological Research* 53, s. 423-29.
- Cartwright N., 1997, „Models: the blueprints for laws”, *Philosophy of Science, Proceedings* 64, s. S292-S303.
- Cartwright N., 1999, *The dappled world. A study of the boundaries of science*, Cambridge: Cambridge University Press.
- Cartwright N., 2001, „What is wrong with bayes nets?”, *Monist* 84, s. 242-64.
- Cartwright N., 2002, „In favor of laws that are not *ceteris paribus* after all”, *Erkenntnis* 57, s. 425-39.
- Cartwright N., 2003, „Two theorems on invariance and causality”, *Philosophy of Science* 70, s. 203-24.
- Cartwright N., 2004, „Causation: one word, many things”, *Philosophy of Science, Proceedings* 72, s. S805-19.
- Cartwright N., Jones M., 1991, „How to hunt quantum causes”, w: W. Spohn (red.), *Erkenntnis orientated: a centennial volume for Rudolf Carnap and Hans Reichenbach*, Dordrecht: Kluwer Academic Publishers, s. 205-31.
- Cheng P., 1997, „From covariation to causation: a causal power theory”, *Psychological Review* 104, s. 367-405.
- Christ C. F., 1960, „A symposium on simultaneous equation estimation: simultaneous equation estimation: any verdict yet?”, *Econometrica* 28, s. 835-45.
- Chu T., Glymour C., Scheines R., Spirtes P., 2003, „A statistical problem for inference to regulatory structure from associations of gene expression measurements with microarrays”, *Bioinformatics* 19, s. 1147-52.
- Churchland P., Hooker C. A. (red.), 1985, *Images of science*, Chicago: Chicago University Press.
- Cichosz P., 2000, *Systemy uczące się*, Warszawa: WNT.
- Cliff N., 1987, „Comments on Professor Freedman's paper”, *Journal of Educational Statistics* 12, s. 158-60.
- Clogg C. C., Haritou A., 1997, „The regression method of causal inference and a dilemma confronting this method”, w: (McKim i Turner 1997), s. 83-112.
- Cochran W. G., 1957, „Analysis of covariance: its nature and uses”, *Biometrics* 13, s. 261-81.
- Cochran W. G., 1968, „The effectiveness of adjustment by subclassification in removing bias in observational studies”, *Biometrics* 24, s. 295-313.
- Cochran W. G., Chambers S. P., 1965, „The planning of observational studies of human populations”, *Journal of the Royal Statistical Society. Series A* 128, s. 234-66.
- Coleman J. S., 1986, „Social theory, social research and a theory of action”, *American Journal of Sociology* 91, s. 1309-35.

- Coleman J. S., 1990, *Foundations of social theory*, Cambridge: Harvard University Press.
- Colwell R. R., 1996, „Global climate and infectious disease: the cholera paradigm”, *Science* 5295, s. 2025-31.
- Collingwood R., 1972, *An essay on metaphysics*, Chicago: Regnery.
- Collins J., Hall N., Paul L. (red.), 2002, *Counterfactuals and causes*, Cambridge, MA: MIT Press.
- Cooper G., 1996, „Theoretical modeling and biological laws”, *Philosophy of Science, Proceedings* 63, s. S28-S35.
- Cooper G. F., 1999, „On overview of the representation and discovery of causal relationships using Bayesian networks”, w: (Glymour i Cooper 1999), s. 3-62.
- Corfield D., Williamson J. (red.), 2001, *Foundations of Bayesianism*, Dordrecht: Kluwer Academic Publishers.
- Cox D. R., 1986, „Statistics and causal inference: comment”, *Journal of the American Statistical Association* 81, s. 963-64.
- Cox D. R., 1990, Role of models in statistical analysis, *Statistical Science* 5, s. 169-74.
- Cox D. R., 1992, Causality: some statistical aspects, *Journal of the Royal Statistical Society, Series A* 155, s. 291-301.
- Cox D. R., Wermuth N., 1995, „Discussion of causal diagrams for empirical research by J. Pearl”, *Biometrika* 82, s. 688-89.
- Cox D. R., Wermuth N., 1996, *Multivariate dependencies: models, analysis and interpretation*, Londyn: Chapman and Hall.
- Cox D. R., Wermuth N., 2001, „Some statistical aspects of causality”, *European Sociological Review* 17, s. 65-74.
- Crámer H., 1946, *Mathematical methods of statistics*, Princeton: Princeton University Press.
- Czarnocka M. (red.), 1995, *Dziedzictwo logicznego empiryzmu*, Warszawa: IFiS PAN.
- Czeżowski T., 1933, *Jak powstało zagadnienie przyczynowości. Zarys jego rozwoju w filozofii starożytnej*, Wilno: WTF.
- Czeżowski T., 1946, *Główne zasady nauk filozoficznych*, Toruń: KN T. Szczęśny i S-ka.
- Czeżowski T., 1946/1958, „O naukach humanistycznych”, w: (Czeżowski 1958), s. 44-58.
- Czeżowski T., 1947/1958, „Twierdzenia ogólne w teorii naukowej”, w: (Czeżowski 1958), s. 105-11.
- Czeżowski T., 1949/1958, „Kilka uwag o racjonalizmie i empiryzmie”, w: (Czeżowski 1958), s. 26-33.
- Czeżowski T., 1958, *Odczyty filozoficzne*, Łódź: PWN.
- Czeżowski T., 1965, *Filozofia na rozdrożu. Analizy metodologiczne*, Warszawa: PWN.
- Czeżowski T., 1968, *Logika. Podręcznik dla studiujących nauki filozoficzne*, Warszawa: PWN.
- D'Agostino R., Rubin D. B., 2000, „Estimation and use of propensity scores with incomplete data”, *Journal of the American Statistical Association* 95, s. 749-59.

- Dale I. A., 1991, *A history of inverse probability: from Thomas Bayes to Karl Pearson*, Nowy Jork: Springer.
- Darnell A. C., 1994, *A dictionary of econometrics*, Brookfield: Elgar.
- Davidson D., 1967/2001, „Causal relations”, w: (Davidson 2001), s. 149-62.
- Davidson D., 2001, *Essays on actions and events: philosophical essays volume 1*, Oxford: Oxford University Press.
- Davis G. C., 2000, „A semantic interpretation of Haavelmo's structure of econometrics”, *Economics and Philosophy* 16, s. 205-28.
- Dawes R. M., 2002, „The ethics of using or not using statistical prediction rules in psychological practice and related consulting activities”, *Philosophy of Science, Proceedings* 69, s. S178-S84.
- Dawid A. P., 1979, „Conditional independence in statistical theory”, *Journal of the Royal Statistical Society, Series B* 41, 1-31.
- Dawid A. P., 1995, „Discussion of causal diagrams for empirical research by J. Pearl”, *Biometrika* 82, s. 689-90.
- Dawid A. P., 2000, „Causality without counterfactuals”, *Journal of the American Statistical Association* 95, s. 407-48.
- Dawid A. P., 2001, „Separoids: a mathematical framework for conditional independence and irrelevance”, *Annals of Mathematics and Artificial Intelligence* 32, s. 335-72.
- De Finetti B., 1969, „Initial probabilities: a prerequisite for any valid induction”, *Synthese* 20, s. 3-17.
- De Finetti B., 1980, „On the condition of partial exchangeability”, w: R. C. Jeffrey (red.), *Studies in inductive logic and probability*, Berkeley: University of California Press, t. 2, s. 193-205.
- De Finetti B., 1982, „Probability: the different views and terminologies in a critical analysis”, w: L. J. Cohen i in. (red.), *Logic, methodology and philosophy of science VI*, Amsterdam: North-Holland, s. 391-94.
- De Finetti B., 1992, *Theory of probability: a critical introductory treatment*, Nowy Jork: Wiley.
- de la Fuente A., Bing N., Hoeschele I., Mendes P., 2004, „Discovery of meaningful associations in genomic data using partial correlation coefficients”, *Bioinformatics* 20, s. 3565-74.
- Dempster A. P., 1991, „Causality and statistics”, *Journal of Statistics Planning and Inference* 25, s. 261-78.
- Dobosz M., 2001, *Wspomagana komputerowo statystyczna analiza wyników badań*, Warszawa: Exit.
- Doll R., 2002, „Proof of causality”, *Perspectives in Biology and Medicine* 45, s. 499-515.
- Doll R., Hill A. B., 1950, „Smoking and carcinoma of the lung”, *British Medical Journal* 4682, s. 739-48.

- Doll R., Hill A. B., 1952, „A study of the aetiology of carcinoma of the lung”, *British Journal of Medicine* 4797, s. 1271-86.
- Doll R., Hill A. B., 1954, „The mortality of doctors in relation to their smoking habits”, *British Medical Journal* 4877, s. 1451-55.
- Doll R., Hill A. B., 1956a, „Lung cancer and other causes of death in relation to smoking”, *British Medical Journal* 5001, s. 1071-81.
- Doll R., Hill A. B., 1956b, „Lung cancer and tobacco”, *British Medical Journal* 4976, s. 1160-63.
- Doll R., Peto R., Boreham J., Sutherland I., 2004, „Mortality in relation to smoking: 50 years' observations on male British doctors”, *British Medical Journal* 7455, s. 1519.
- Dorling J., 1992, „Bayesian conditionalization resolves positivist/realist disputes”, *The Journal of Philosophy* 89, s. 362-82.
- Dowe P., 2000, *Physical causation*, Cambridge: Cambridge University Press.
- Dowe P., 2005, „Causation and misconceptions”, *Philosophy of Science, Proceedings* 72, s. S926-31.
- Dretske F. I., 1977, „Laws of nature”, *Philosophy of Science* 44, s. 248-68.
- Duncan D., 1966, „Path analysis: sociological examples”, *The American Journal of Sociology* 72, s. 1-16.
- Duncan O. D., 1975, *Introduction to structural equation models*, Nowy Jork: Academic Press.
- Earman J., 1978, „The universality of laws”, *Philosophy of Science* 45, s. 173-81.
- Earman J., 1992, *Bayes or bust? A critical examination of Bayesian confirmation theory*, Cambridge, MA: MIT Press.
- Earman J. (red.), 1992, *Inference, explanation, and other frustrations: essays in the philosophy of science*, Berkeley: University of California Press.
- Earman J., 1993, „In defense of laws: reflections on Bas van Fraassen's *Laws and symmetry*”, *Philosophy and Phenomenological Research* 53, s. 413-19.
- Earman J., Roberts J., 1999, „*Ceteris paribus*, there is no problem of provisos”, *Synthese* 118, s. 439-78.
- Earman J., Roberts J., Smith S., 2002, „*Ceteris paribus* lost”, *Erkenntnis* 57, s. 281-301.
- Edelman E., 2001, „On causes: Hippocrates, Aristotle, Robert Koch, and the Dread Pirate Roberts”, *Circulation* 104, s. 2509-12.
- Eells E., 1991, *Probabilistic causality*, Cambridge: Cambridge University Press.
- Efron B., Feldman D., 1991, „Compliance as an explanatory variable in clinical trials”, *Journal of the American Statistical Association* 86, s. 9-26.
- Enders W., 1989, *Applied econometric time series*, Nowy Jork: Wiley.
- Engle R. F., Hendry D., Richard J. F., 1983, „Exogeneity”, *Econometrica* 51, s. 277-304.
- Feyerabend P., 1962/1979, „Wyjaśnianie, redukcja i empiryzm”, w: (Feyerabend 1979), s. 62-151.

- Feyerabend P., 1964/1979, „Realizm i instrumentalizm. Uwagi o logice potwierdzania przez fakty”, w: (Feyerabend 1979), s. 152-93.
- Feyerabend P., 1969/1979, „Nauka bez doświadczenia”, w: (Feyerabend 1979), s. 194-99.
- Feyerabend P., 1979, *Jak być dobrym empirystą?*, tłum. i wstęp K. Zamiara, Warszawa: PWN.
- Fellows P., Linscott M., Ivins B., Pitt M., Rossi C., Gibbs P., Friedlander A., 2001, „Efficacy of a human anthrax vaccine in guinea pigs, rabbits, and rhesus macaques against challenge by *Bacillus anthracis* isolates of diverse geographical origin”, *Vaccine* 19, s. 3241-47.
- Ferguson G. A., Takane Y., 1989/2003, *Analiza statystyczna w psychologii i pedagogice*, tłum. M. Zagrodzki, Warszawa: PWN.
- Fienberg S., Glymour C., Spirtes P., 1995, „Discussion of causal diagrams for empirical research by J. Pearl”, *Biometrika* 82, s. 690-92.
- Fisher R. A., 1935, *The design of experiments*, Edynburg: Oliver and Boyd.
- Fisher R. A., 1959, *Smoking: the cancer controversy. Some attempts to assess the evidence*, Edynburg: Oliver and Boyd.
- Fisher R. A., 1973, *Statistical methods and scientific inference*, Nowy Jork: Hafner.
- Fox J., 1987, „Statistical models for non-experimental data: a comment on Freedman”, *Journal of Educational Statistics* 12, s. 161-64.
- Frangakis C., Rubin D. B., 2002, „Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes”, *Biometrika* 86, s. 365-79.
- Frank K., 2000, „Impact of a confounding variable on a regression coefficient”, *Sociological Methods and Research* 29, s. 147-94.
- Freedman D. A., 1983, *Structural-equation models: a case study*, Berkeley: University of California Press, *Technical report No. 22. Department of Statistics*.
- Freedman D. A., 1985, *Statistical analysis and shoe leather*, *Sociological Methodology* 21, s. 291-313.
- Freedman D. A., 1987a, „A rejoinder on models, metaphors and fables”, *Journal of Educational Statistics* 12, s. 206-23.
- Freedman D. A., 1987b, „As others see us: a case study in path analysis”, *Journal of Educational Statistics* 12, s. 101-28.
- Freedman D. A., 1995, „Discussion of causal diagrams for empirical research by J. Pearl”, *Biometrika* 82, s. 692-93.
- Freedman D. A., 1997a, „From association to causation via regression”, w: (McKim i Turner 1997), s. 113-61.
- Freedman D. A., 1997b, „Rejoinder to Spirtes and Scheines”, w: (McKim i Turner 1997), s. 176-182.
- Freedman D. A., 1999, „From association to causation: some remarks on the history of statistics”, *Statistical Science* 14, s. 243-58.

- Freedman D. A., 2002, *On specifying graphical models for causation, and the identification problem*, Department of Statistics, University of California, Berkeley.
- Freedman D. A., Humphreys P., 1999, „Are there algorithms that discover causal structure?”, *Synthése* 121, s. 29-54.
- Freedman D. A., Petitti D. B., 2002, „Salt, blood pressure, and public policy”, *International Journal of Epidemiology* 31, s. 319-20 .
- Freedman D. A., Zeisel H., 1988a, „From mouse-to-man: the quantitative assessment of cancer risks”, *Statistical Science* 3, s. 3-28.
- Freedman D. A., Zeisel H., 1988b, „From mouse-to-man: the quantitative assessment of cancer risks - rejoinder”, *Statistical Science* 3, s. 45-56.
- Friedman M., 1974, „Explanation and scientific understanding”, *The Journal of Philosophy* 71, s. 5-19.
- Friedman M., 1981, „Theoretical explanation”, w: R. Healey (red.), *Reduction, time, and reality: studies in the philosophy of the natural sciences*, Cambridge: Cambridge University Press, s. 1-16.
- Frisch R., 1939/1995, „Autonomy of economic relations”, w: (Hendry i Morgan 1995), s. 407-19.
- Gatnar E., 2002, „Statystyczne modele związków przyczynowo-skutkowych”, w: (Mynarski 2002), s. 95-103.
- Gaul M., Machowski A., 1987, „Elementy analizy ścieżek”, J. Brzeziński (red.), *Wielozmiennowe modele statystyczne w badaniach psychologicznych*, Warszawa: PWN, s. 82-112.
- Gawecki B., 1969, *Zagadnienie przyczynowości w fizyce*, Warszawa: IW Pax.
- Geiger D., Verma T., Pearl J., 1990, „Identifying independence in Bayesian networks”, *Networks* 20, s. 507-34.
- Gelman S., 2004, „Psychological essentialism in children”, *Trends in Cognitive Sciences* 8, s. 404-409.
- Giedymin J., 1964, *Problemy, założenia, rozstrzygnięcia. Studia nad logicznymi podstawami nauk społecznych*, Warszawa: PWN.
- Giedymin J., 1975, „Antipositivism in contemporary philosophy of social science”, *Britisch Journal for the Philosophy of Science* 26, s. 275-301.
- Giere R. N., 1985, „Philosophy of science naturalized”, *Philosophy of Science* 52, s. 331-56.
- Gillies D., 2002, „Causality, propensity, and Bayesian networks”, *Synthése* 132, s. 63-88.
- Giza P., 1990, *Realizm Iana Hackinga a konstruktywny empiryzm Bas C. van Fraassena*, Lublin: W UMCS.
- Glymour C., 1980, *Theory and evidence*, Princeton: Princenton University Press.
- Glymour C., 1985, „Explanation and realism”, w: P Churchland, C. A. Hooker (red.), *Images of science*, Chicago: Chicago University Press, s. 99-117.
- Glymour C., 1986, „Statistics and causal inference: comment: statistics and metaphysics”, *Journal of the American Statistical Association* 81, s. 964-66.

- Glymour C., 2001, *The mind's arrows: Bayes nets and graphical causal models in psychology*, Cambridge, MA: MIT Press/A Bradford Book.
- Glymour C., 2004, „The automation of discovery”, *Daedalus* 133, s. 69-77.
- Glymour C., Cooper G. F. (red.), 1999, *Computation, causation, and discovery*, Cambridge, MA: MIT Press.
- Glymour C., Kelly K. T., 2004, „Why probability does not capture the logic of scientific justification”, w: (Hitchcock 2004), s. 94-114.
- Glymour C., Scheines R., Spirtes P., 1988, „Exploring causal structure with the TETRAD program”, *Sociological Methodology* 18, s. 411-48.
- Glymour C., Spirtes P., Scheines R., 1991, „Causal inference”, w: (Spohn 1991), s. 151-89.
- Glymour C., Spirtes P., 1988, „Latent variables, causal models and overidentifying constraints”, *Journal of Econometrics* 39, s. 175-98.
- Glymour C., Scheines R., Spirtes P., Kelly K., 1987, *Discovering causal structure: artificial intelligence, philosophy of science, and statistical modeling*, Nowy Jork: Academic Press.
- Glymour C., Spirtes P., Richardson T., 1999, „On the possibility of inferring causation from association without background knowledge”, w: (Glymour i Cooper), s. 323-31.
- Goldberger A. S., 1972, „Structural equation methods in the social sciences”, *Econometrica* 40, s. 979-1001.
- Goldthorpe J. H., 2000, *On sociology: numbers, narratives, and the integration of research and theory*, Oxford: Oxford University Press.
- Goldthorpe J. H., 2001, „Causation, statistics, and sociology”, *European Sociological Review* 17, s. 1-20.
- Good I. J., 1965, *The estimation of probabilities: an essay on modern Bayesian methods*, Cambridge, MA: MIT Press.
- Goodman N., 1947/1948, „On infirmities of confirmation-theory”, *Philosophy and Phenomenological Research* 8, s. 149-51.
- Goodman N., 1983, *Fact, fiction, and forecast*, Cambridge: Harvard University Press.
- Gopnik A., Glymour C., Sobel D., Schulz L., Kushnir T., Danks D., 2004, „A theory of causal learning in children: causal maps and Bayes nets”, *Psychological Review* 111, s. 3-32.
- Gopnik A., Schulz L., 2004, „Mechanisms of theory formation in young children”, *Trends in Cognitive Sciences* 8, s. 371-77.
- Gopnik A., Sobel D., Schulz L., Glymour C., 2001, „Causal learning mechanisms in very young children: two-, three-, and four-year-olds infer causal relations from patterns of variation and covariation”, *Developmental Psychology* 37, s. 620-29.
- Gordon A. D., Vichi M., 1998, „Partitions of partitions”, *Journal of Classification* 15, s. 265-85.
- Grabski F., Jaźwiński J., 2001, *Metody bayesowskie w niezawodności i diagnostyce*, Warszawa: WKŁ.

- Granger C., 1969, „Investigating causal relations by econometric models and cross-spectral methods”, *Econometrica* 37, s. 424-38.
- Granger C., 1986, „Statistics and causal inference: comment”, *Journal of the American Statistical Association* 81, s. 967-68.
- Granger C., 2003, „Some aspects of causal relationships”, *Journal of Econometrics* 112, s. 69-71.
- Granger C., 2005, „Comments on the 20th anniversary issue of *Econometric Theory*”, *Econometric Theory* 21, s. 298.
- Granger C., Hendry D. F., 2005, „A dialogue concerning a new instrument for econometric modeling”, *Econometric Theory* 21, s. 278-97.
- Granger C., Jeon Y., 2003, „Interactions between large macro models and time series analysis”, *International Journal of Finance & Economics* 8, s. 1-10.
- Greenland S., Poole C., 1988, „Invariants and noninvariants in the concept of interdependent effects”, *Scandinavian Journal of Work and Environmental Health* 14, s. 125-29.
- Grobler A., 1993, *Prawda i racjonalność naukowa*, Kraków: Interesse.
- Grobler A., 1994, „Justification of the empirical basis: the Popperian vs. the inductivist conception of rationality”, w: G. Meggle, U. Wessles (red.), *Analyomen I. Proceedings of the 1 conference 'Perspectives in analytical philosophy'*, Berlin: de Gruyter, s. 299-309.
- Gutenbaum J., 1987, *Modelowanie matematyczne systemów*, Warszawa: PWN.
- Guilford J. P., 1954/2005, „Ograniczenia regresji wielokrotnej”, tłum. E. Hornowska, M. Zakrzewska, w: J. Brzeziński (red.), *Trafność i rzetelność testów psychologicznych. Wybór tekstów*, Gdańsk: GWP, s. 97-101.
- Gumański L., 2004, „Uwagi o związku przyczynowym”, *Ruch Filozoficzny* 61, s. 109-15.
- Haavelmo T., 1943, „The statistical implications of a system of simultaneous equations”, *Econometrica* 11, s. 1-12.
- Haavelmo T., 1944, „The probability approach in econometrics”, *Econometrica* 12, Supplement s. 1-115.
- Hacking I., 1965, *The logic of statistical inference*, Cambridge: Cambridge University Press.
- Hacking I., 1984, *Representing and intervening*, Cambridge: Cambridge University Press.
- Hajduk Z., 1970, „Wyjaśnianie dedukcyjne”, *Roczniki Filozoficzne* 18, z. 3, s. 69-99.
- Hajduk Z., 1977, „Rodzaje empiryzmu w metodologii”, *Summariusz* 26, s. 209-15.
- Hajduk Z., 2002, „Geneza teorii naukowych: uaktualniona forma kontrowersji o pluralizm metodologii”, *Roczniki Filozoficzne* 50, z. 3.
- Hall N., 2005, „Rescued from the rubbish bin: Lewis on causation”, *Philosophy of Science, Proceedings* 72, s. S1107-14.
- Hankinson R. J., 1998, *Cause and explanation in Ancient Greek thought*, Oxford: Clarendon Press.
- Hansen B., 2005, „Challenges for econometric model selection”, *Econometric Theory* 21, s. 60-68.

- Harris T., 2003, „Data models and the acquisition and manipulation of data”, *Philosophy of Science* 70, s. 1508-17.
- Hausman D. M., 1997, „Causation, agency, and independence”, *Philosophy of Science* 64 Supplement, s. S15-S25.
- Hausman D. M., 1998, *Causal asymmetries*, Cambridge: Cambridge University Press.
- Hausman D. M., 1999, „The mathematical theory of causation”, *British Journal for the Philosophy of Science* 50, s. 151-62.
- Hausman D. M., Woodward J., 1999, „Independence, invariance and the causal Markov condition”, *British Journal for the Philosophy of Science* 50, s. 521-83.
- Hausman D. M., Woodward J., 2004, „Modularity and the causal Markov condition: a restatement”, *The British Journal for the Philosophy of Science* 55, s. 147-61.
- Hausman D. M., Woodward J., 2005, „Manipulation and the Causal Markov Condition”, *Philosophy of Science, Proceedings* 72, s. S846-56.
- Hausman J., 2003, „Triangular structural model specification and estimation with application to causality”, *Journal of Econometrics* 112, s. 107-13.
- Heckman J. J., 1989, „Causal inference and nonrandom samples”, *Journal of Educational Statistics* 14, s. 159-68.
- Heckman J. J., 1996, „Identification of causal effects using instrumental variables: comment”, *Journal of the American Statistical Association* 91, s. 459-62.
- Heckman J. J., 2000, „Causal parameters and policy analysis in economics: a twentieth century retrospective”, *The Quarterly Journal of Economics* 115, s. 45-97.
- Heckman J. J., 2003, „Conditioning, causality and policy analysis”, *Journal of Econometrics* 112, s. 73-78.
- Heller M., 2002, *Czas i przyczynowość*, Lublin: TN KUL.
- Heller M., Życiński J., 1987, „Epistemologiczne aspekty związków filozofii z nauką”, w: M. Heller, A. Michalik, J. Życiński (red.), *Filozofować w kontekście nauki*, Kraków: PTT, s. 7-17.
- Hempel C. G., 1965, *Aspects of scientific explanation*, Nowy Jork: The Free Press.
- Hempel C. G., Oppenheim P., 1948, „Studies in the logic of explanation”, *Philosophy of Science* 15, s. 135-75.
- Hendricks V. F., 2001, *The convergence of scientific knowledge: a view from the limit*, Dordrecht: Kluwer Academic Publishers.
- Hendry D., 1995, *Dynamic econometrics*, Nowy Jork: Oxford University Press.
- Hendry D. F., 2001, „Achievements and challenges in econometric methodology”, *Journal of Econometrics* 100, s. 7-10.
- Hendry D. F., Krolzig H. M., 2004, „Automatic model selection: a new instrument for social science”, *Electoral studies* 23, s. 525-44.
- Hendry D. F., Mizon G., 2000, „Reformulation of empirical macroeconomic modelling”, *Oxford Review of Economic Policy* 16, s. 138-59.

- Hendry D., Morgan M., 1995, „Introduction”, w: D. Hendry, M. Morgan (red.), *The foundations of econometric analysis*, Cambridge: Cambridge University Press, 1-82.
- Hershberger S. L., 2003, „The growth of structural equation modeling: 1994-2001”, *Structural Equation Modeling* 10, s. 35-46.
- Hildreth C., 1960, „Simultaneous equation estimation: any verdict yet?”, *Econometrica* 28, s. 846-54.
- Hill B. A., 1937/1991, *Principles of medical statistics*, Londyn: Arnold.
- Hill B. A., 1965, „The environment and disease: association or causation”, *Proceedings of the Royal Society of Medicine* 58, s. 295-300.
- Hitchcock C., 2001a, „A tale of two effects”, *The Philosophical Review* 110, 361-96.
- Hitchcock C., 2001b, „Causal generalizations and good advice”, *Monist* 84, s. 218-41.
- Hitchcock C., 2003a, „Of Humean bondage”, *The British Journal for the Philosophy of Science* 54, s. 1-25.
- Hitchcock C., 2003b, „Unity and plurality in the concept of causation”, w: F. Stadler (red.), *The Vienna Circle and logical empiricism: re-evaluation and future perspectives*, Dordrecht: Kluwer Academic Publishers, s. 217-24.
- Hitchcock C. (red.), 2004, *Contemporary debates in philosophy of science*, Oxford: Basil Blackwell.
- Hitchcock C., 2005, „Causal processes and interactions: what are they and what are they good for?”, *Philosophy of Science, Proceedings* 72, s. S932-41.
- Hofer-Szabó G., Rédei M., Szabó L., 1999, „On Reichenbach's common cause principle and Reichenbach's notion of common cause”, *British Journal for the Philosophy of Science* 50, s. 377-99.
- Holland P., 1986a, „Rejoinder”, *Journal of the American Statistical Association* 81, s. 968-70.
- Holland P., 1986b, „Statistics and causal inference”, *Journal of the American Statistical Association* 81, s. 945-60.
- Holland P., 1988, „Causal inference, path analysis, and recursive structural equation models”, *Sociological Methodology* 18, s. 449-84.
- Hoover K. D., 2001, *Causality in macroeconomics*, Cambridge: Cambridge University Press.
- Hoover K. D., 2003, „Some causal lessons from macroeconomics”, *Journal of Econometrics* 112, s. 121-25.
- Hoover K. D., 2004, „Lost causes”, *Journal of the History of Economic Thought* 26, s. 149-64.
- Hoover K. D., 2005, „Automatic inference of the contemporaneous causal order of a system of equations”, *Econometric Theory* 21, s. 69-77.
- Hope K., 1987, „Barren theory or petty craft? A response to Professor Freedman”, *Journal of Educational Statistics* 12, s. 129-47.
- Horwich P., 1982, *Probability and evidence*, Cambridge: Cambridge University Press.
- Howard-Jones N., 1984, „Robert Koch and the cholera vibrio: a centenary”, *British Medical Journal* 288, s. 379-81.

- Howarth R. J., 2001, „A history of regression and related model-fitting in the earth sciences (1636?-2000)”, *Natural Resources Research* 10, s. 241-86.
- Howie D., 2002, *Interpreting probability: controversies and developments in the early twentieth century*, Cambridge: Cambridge University Press.
- Howson C., Urbach P., 1993, *Scientific reasoning: the Bayesian approach*, Illinois: Open Court.
- Hume D., 1939/1963, *Traktat o naturze ludzkiej*, t. 1, tłum. C. Znamierowski, Warszawa: PWN.
- Hume D., 1748/1977, *Badania dotyczące rozumu ludzkiego*, tłum. J. Łukasiewicz, K. Twardowski, oprac. A. Hochfeldowa, Warszawa: PWN.
- Humphreys P., 1995, „Computational empiricism”, *Foundations of Science* 1, s. 119-30.
- Humphreys P., 2005, „Some thoughts on Wesley Salmon's contributions to the philosophy of probability”, *Philosophy of Science, Proceedings* 72, s. S942-49.
- Hurwicz L., 1944, „Stochastic models of economic fluctuations”, *Econometrica* 12, s. 114-24.
- Imai K., van Dyk D. A., 2004, „Causal inference with general treatment regimes: generalizing the propensity score”, *Journal of the American Statistical Association* 99, s. 854-66.
- Imbens G. W., Rubin D. B., 1995, „Discussion of causal diagrams for empirical research by J. Pearl”, *Biometrika* 82, s. 694-95.
- Ingarden R., 1945, „Quelques remarques sur la relation de causalité”, *Studia Philosophica* 3, s. 151-66.
- Ingarden R., 1955, „Niektóre twierdzenia o związku przyczynowym”, *Sprawozdania Towarzystwa Naukowego w Toruniu* 9, s. 77-82. Autorem sprawozdania był Tadeusz Czeżowski.
- Jadacki J. J., 2005, „O związku przyczynowo-skutkowym”, *Ruch Filozoficzny* 62, s. 217-22.
- Jadamus-Hacura M., 2002, „Bayesowska analiza modeli autoregresyjnych i średniej ruchomej”, *Badania Operacyjne i Decyzje* 1, s. 33-40.
- Jahoda M., Deutsch M., Cook S. W., 1965, „Ustalanie związków przyczynowych w badaniach nieeksperymentalnych”, w: Nowak S. (red.), *Metody badań socjologicznych*, Warszawa: PWN, s. 561-71.
- Jain S., Osherson D., Royer J., Sharma A., 1999, *Systems that learn: an introduction to learning theory*, Cambridge, MA: MIT Press/A Bradford Book.
- Janiszowski K., 2002, *Identyfikacja modeli parametrycznych*, Warszawa: Exit.
- Jeffrey R. C., 1966, „Goodman's query”, *The Journal of Philosophy* 63, s. 281-88.
- Jeffrey R. C., 1969, „Statistical explanation vs. statistical inference”, w: N. Rescher (red.), *Essays in honor of Carl G. Hempel*, Dordrecht: Reidel, s. 104-113.
- Jeffrey R. C., 1995, „A brief guide to the work of Carl Gustav Hempel”, *Erkenntnis* 42, s. 3-7.
- Jensen F. V., 1996, *An introduction to Bayesian networks*, Nowy Jork: Springer.
- Judycki S., 1995, *Umysł i synteza*, Lublin: RW KUL.

- Judycki S., 2005, „Ewolucja epistemologii w XX wieku: kontynuacja czy zmiana tematu?”, w: R. i I. Ziemiński (red.), *Byt i sens. Księga pamiątkowa VII Polskiego Zjazdu Filozoficznego w Szczecinie*, Szczecin: WN US s. 140-48.
- Juhl C., 1993, „Bayesianism and reliable scientific inquiry”, *Philosophy of Science* 60, s. 302-19.
- Juhl C., Kelly K., 1994, „Realism, convergence, and additivity”, *Philosophy of Science, Proceedings* 61, s. 181-89.
- Kaluszyńska E., 1994, „Czego nie wolno van Fraassenowi, czyli o tożsamości filozofii nauki”, *Filozofia Nauki* 7-8, s. 181-87.
- Kaluszyńska E., 1995, *Modele teorii empirycznych*, Warszawa: IFiS PAN.
- Kaluszyńska E., 2001, „Realizm w filozofii nauki”, *Analecta Cracoviensia* 33, s. 71-93.
- Kamiński S., 1966/1989, „Wyjaśnianie w metafizyce”, w: S. Kamiński, *Jak filozofować? Studia z metodologii filozofii klasycznej*, oprac. T. Szubka, Lublin: TN KUL, s. 151-76.
- Kamiński S., 1976/1994, „Metodologia nauk. Współczesne problemy i tendencje”, w: (Kamiński 1994), s. 419-33.
- Kamiński S., 1981/1994, „O klasyfikacji rozumowań”, w: (Kamiński 1994), s. 231-45.
- Kamiński S., 1992, *Nauka i metoda. Pojęcie nauki i klasyfikacja nauk*, oprac. A. Bronk, Lublin: TN KUL.
- Kamiński S., 1994, *Metoda i język. Studia z semiotyki i metodologii nauk*, oprac. U. Żegleń, Lublin: TN KUL.
- Kamlah A., 1991, „The causal relation as the most fundamental fact of the world. Comments on Hans Reichenbach's paper: The space problem in the new quantum mechanics”, w: (Spohn 1991), s. 49-60.
- Kaprio J., Koskenvuo M., 1989, „Twins, smoking and mortality: a 12-year prospective study of smoking-discordant twin pairs”, *Social Science and Medicine* 29, s. 1083-89.
- Karhausen L. R., 2000, „Causation: the elusive grail of epidemiology”, *Medicine Health Care and Philosophy* 3, 59-67.
- Karlin S., 1987, „Path analysis in genetic epidemiology and alternatives”, *Journal of Educational Statistics* 12, s. 165-77.
- Karpiński J., 1985, *Przyczynowość w badaniach socjologicznych*, Warszawa: PWN.
- Kaufman J. S., Poole C., 2000, „Looking back on *Causal thinking in the health sciences*”, *Annual Reviews of Public Health* 21, s. 101-19.
- Kawalec P., 2001, „Back to green perspectives on confirmation as justification”, http://www.jtb-forum.pl/jtb/papers/pk_btgpocaj.pdf.
- Kawalec P., 2003a, *Structural reliabilism: inductive logic as a theory of justification*, Dordrecht: Kluwer Academic Publishers.
- Kawalec P., 2003b, „Zagadnienia metodologiczne w bayesowskiej teorii konfirmacji”, *Roczniki Filozoficzne* 51, z. 1, s. 113-42.
- Kawalec P., 2004, „Wyjaśnić to podać model przyczynowy”, *Roczniki Filozoficzne* 52, s. 241-66.

- Kawalec P., 2005, „Understanding science for the new millenium”, *Archives for the Philosophy of Science*; <http://philsci-archive.pitt.edu/archive/00002558>.
- Kelly K. T., 1996, *The logic of reliable inquiry*, Oxford: Oxford University Press.
- Kelly K. T., 2000, „The logic of success”, *British Journal for the Philosophy of Science* 51, s. 639-66.
- Kiczuk S., 1995, *Związek przyczynowy a logika przyczynowości*, Lublin: RW KUL.
- Kiiveri H., Speed T., 1982, „Structural analysis of multivariate data: a review”, *Sociological Methodology* 13, s. 209-89.
- Kim J., 1999, „Hempel, explanation, metaphysics”, *Philosophical Studies* 94, s. 1-20.
- Kincaid H., 2004, „There are laws in the social sciences”, w: (Hitchcock 2004), s. 168-185.
- Kistler M., 2002, „The necessity of laws, quiddity and the causal criterion of reality”, w: A. Beckermann, C. Nitz (red.), *Argument und Analyse*, Paderborn: Mentis, s. 461-73.
- Kitcher P., 1976, „Explanation, conjunction, and unification”, *The Journal of Philosophy* 73, s. 207-12.
- Kitcher P., 1981, „Explanatory unification”, *Philosophy of Science* 48, s. 507-31.
- Kitcher P., 1992, „The naturalists return”, *The Philosophical Review* 101, s. 53-114.
- Kitcher P., 1993, *The advancement of science*, Oxford: Oxford University Press.
- Kitcher P., 2001, „Real realism: the Galilean strategy”, *The Philosophical Review* 110, s. 151-97.
- Kitcher P., Salmon W.C., 1998, „Van Fraassen on explanation”, w: W. C. Salmon, *Causality and explanation*, Nowy York: Oxford University Press, s. 178-90.
- Klee R. (red.), 1999, *Scientific inquiry: readings in the philosophy of science*, Oxford: Oxford University Press.
- Klein L. R., 1960, „Single equation vs. equation system methods of estimation in econometrics”, *Econometrica* 28, s. 866-71.
- Kłopotek M. A., Wierzchoń S. T., 2002, *Evidential reasoning. An interpretative investigation*, Siedlce: Akademia Podlaska.
- Kmita J., 1977, *Wykłady z logiki i metodologii nauk*, Warszawa: PWN.
- Kmita J., Nowak L., 1968, *Studia nad teoretycznymi podstawami humanistyki*, Poznań: WN UAM.
- Kneale W., 1949, *Probability and induction*, Oxford: Clarendon.
- Koopmans T. C., 1937, *Linear regression analysis of economic time series*, Haarlem: Netherlands Economic Institute.
- Koopmans T. C., 1947, „Measurement without theory”, *The Review of Economic Statistics* 29, s. 161-72.
- Koopmans T. C., 1949, „Identification problems in economic model construction”, *Econometrica* 17, s. 125-44.

- Korb K., Wallace C., 1997, „In search of the philosopher's stone: remarks on the Humphreys and Freedman's critique of causal discovery”, *British Journal for the Philosophy of Science* 48, s. 543-53.
- Kripke S., 1972/1988, *Nazywanie a konieczność*, tłum. B. Chwedeńczuk, Warszawa: IW Pax.
- Kotarbiński T., 1961/1986, *Elementy teorii poznania, logiki formalnej i metodologii nauk*, Warszawa: PWN.
- Kotarbiński T., 1965, *Traktat o dobrej robocie*, Warszawa: PWN.
- Krajewski W., 1967, *Związek przyczynowy*, Warszawa: PWN.
- Krajewski W., 1976, *Konieczność, przypadek, prawo statystyczne*, Warszawa: PWN.
- Krajewski W., 1998, *Prawa nauki. Przegląd zagadnień metodologicznych i filozoficznych*, Warszawa: KiW.
- Kuersteiner G., 2005, „Automatic inference for infinite order vector autoregressions”, *Econometric Theory* 21, s. 85-115.
- Kwiatkowski T., 1969, *Poznanie naukowe u Arystotelesa. Niektóre poglądy teoretyczne*, Warszawa: PWN.
- Lad F., 1999, „Assessing the foundation for Bayesian networks: a challenge to the principles and the practice”, *Soft Computing* 3, s. 175-82.
- Lagnado D., Sloman S., 2004, „The advantage of timely intervention”, *Journal of Experimental Psychology: Learning Memory and Cognition* 30, s. 856-76.
- Lalonde R., 1986, „Evaluating the econometric evaluations of training programs”, *American Economic Review* 76, s. 604-20.
- Langley P., Simon H. A., Bradshaw G. L., Żytkow J., 1987, *Scientific discovery. Computational explorations of the creative processes*, Cambridge, MA: MIT Press.
- Lauritzen S. L., 2001, „Causal inference from graphical models”, w: O. E. Barndor-Nielsen, D. R. Cox, C. Klüppelberg (red.), *Complex Stochastic Systems*, Londyn: Chapman and Hall/CRC, s. 63-107.
- Lauritzen S. L., 2004, „Discussion on causality”, *Scandinavian Journal of Statistics* 31, s. 189-201.
- Lazarsfeld P. F., 1961/1968, *Algebra systemów dychotomicznych. Zastosowania matematyki w badaniach socjologicznych*, tłum. i oprac. T. Pawłowski, Warszawa: PWN.
- Lazarsfeld P., 1965, „Ustalanie związków przyczynowych w badaniach nieeksperymentalnych”, w: Nowak S. (red.), *Metody badań socjologicznych*, Warszawa: PWN, s. 571-89.
- Leamer E., 1985, „Vector autoregressions for causal inference?”, *Carnegie-Rochester Conference Series on Public Policy* 22, s. 255-304.
- Leeb H., Pötscher B. M., 2005, „Model selection and inference: facts and fiction”, *Econometric Theory* 21, s. 21-59.
- Leplin J., 1997, *A novel defense of scientific realism*, Oxford: Oxford University Press.
- Lewis D., 1973, *Counterfactuals*, Oxford: Basil Blackwell.

- Lewis D., 1986a, „Causal explanation”, w: D. Lewis, *Philosophical papers*, Nowy Jork: Oxford University Press, s. 214-40.
- Lewis D., 1986b, „Causation”, w: D. Lewis, *Philosophical papers*, Nowy Jork: Oxford University Press, s. 159-213.
- Lewis D., 2000, „Causation as influence”, *The Journal of Philosophy* 97, s. 182-97.
- Lieberson S., 1985, *Making it count*, Berkeley: University of California Press.
- Lindley D. V., 1965, *Introduction to probability and statistics*, Cambridge: Cambridge University Press.
- Little D., 1996, „Causal explanation in the social sciences”, *Southern Journal of Philosophy* 34, Suppl., s. 31-56.
- Little R., 1985, „A note about models for selectivity bias”, *Econometrica* 53, s. 1469-74.
- Liu T. C., 1960, „Underidentification, structural estimation, and forecasting”, *Econometrica* 28, s. 855-65.
- Levi I., 1985, „Common causes, smoking, and lung cancer”, w: R. Campbell, L. Sowden (red.), *Paradoxes of rationality and cooperation*, Vancouver: University of British Columbia Press, s. 234-47.
- Losee J., 1993/2001, *Wprowadzenie do filozofii nauki*, tłum. T. Bigaj, Warszawa: Prószyński.
- Lubański M., 1997, „Informacja – system”, w: M. Heller, M. Lubański, S. W. Ślaga (red.), *Zagadnienia filozoficzne współczesnej nauki*, Warszawa: WATK, s. 13-153.
- Lunceford J. K., Davidian M., 2004, „Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study”, *Statistics in Medicine* 23, s. 2937-60.
- Lukasiewicz J., 1907.1961, „Analiza i konstrukcja pojęcia przyczyny”, w: J. Łukasiewicz, 1961, *Z zagadnień logiki i filozofii. Pisma wybrane*, wybór, wstęp i oprac. J. Śłupecki, Warszawa: PWN 1961, s. 9-62.
- Mackie J. L., 1974, *The cement of the universe: a study of causation*, Oxford: Clarendon.
- Maddala G. S., 1992, *Introduction to econometrics*, Nowy Jork: Macmillan.
- Majdański S., 1972, *Problemy asercji zdaniowej. Szkice pragmatyczne*, Lublin: TN KUL.
- Malewski A., Topolski J., 1960, *Studia z metodologii historii*, Warszawa: PWN.
- Marciszewski W., 1963, „Dzieła sztuki jako znaki”, *Estetyka* 4, s. 231-40.
- Marciszewski W., 1972, *Podstawy logicznej teorii przekonań*, Warszawa: PWN.
- Marciszewski W., 2004, „Nierozstrzygalność i algorytmiczna niedostępność w naukach społecznych”, *Filozofia Nauki* 47-48, s. 5-31.
- Marshall A., 1961, *Principles of economics*, Londyn: Macmillan.
- Mazierski S., 1995, *Prawa przyrody*, Lublin: RW KUL.
- Mączka J., 2001, „Strukturalizm Joachima Metallmanna”, *Analecta Cracoviensia* 33, s. 139-47.

- McKim V. R., Turner S. P. (red.), 1997, *Causality in crisis? Statistical methods and the search for causal knowledge in the social sciences*, Notre Dame: University of Notre Dame Press.
- Mealli F., Rubin D. B., 2002, „Assumptions when analyzing randomized experiments with noncompliance and missing outcomes”, *Health Services and Outcomes Research Methodology* 3, s. 225-32.
- Mealli F., Rubin D. B., 2003, „Assumptions allowing the estimation of direct causal effects”, *Journal of Econometrics* 112, s. 79-87.
- Meek C., 1995, „Causal inference and causal explanation with background knowledge”, w: P. Besnard i S. Hanks (red.), *Uncertainty in artificial intelligence*, t. 11, San Francisco: M. Kaufmann, s. 403-10.
- Meek C., Glymour C., 1994, „Conditioning and intervening”, *British Journal for the Philosophy of Science* 45, s. 1001-21.
- Mejbaum W., 1985, „Powiedzmy: ‘instrumentalizm’”, *Studia Filozoficzne* 5-6, s. 35-53.
- Mejbaum W., 1995, *Wyjaśnianie i wyjaśnienie. Zarys teorii eksplanacji*, Szczecin: FRNF.
- Mellor D. H., 1991, *Matters of metaphysics*, Cambridge: Cambridge University Press.
- Mellor D. H., 1995, *The facts of causation*, Londyn: Routledge.
- Menzies P., 2005, „Causal models, token causation, and processes”, *Philosophy of Science, Proceedings* 72, s. S820-32.
- Menzies P., Price H., 1993, „Causation as a secondary quality”, *British Journal for the Philosophy of Science* 44, s. 187-203.
- Metallmann J., 1929/2002, *Determinizm w biologii*, oprac. i wstęp J. Mączka, Kraków: OBI.
- Metallmann J., 1934, *Determinizm nauk przyrodniczych*, Kraków: PAU.
- Mill J. S., 1843/1962, *System logiki*, t. 1-2, tłum. C. Znamierowski, Warszawa: PWN.
- Mochon D., Sloman S., 2004, „Causal models frame interpretation of mathematical equations”, *Psychonomic Bulletin & Review* 11, s. 1099-1104.
- Mokrzycki E., 1980, *Filozofia nauki a socjologia. Od doktryny metodologicznej do praktyki badawczej*, Warszawa: PWN.
- Morabia A., 2001, „Snow and Farr: a scientific duet”, *Soziologie und Präventivmedizin* 46, 217-24.
- Morgan M. S., Morrison M. (red.), 1999, *Models as mediators: Perspectives on natural and social science*, Cambridge: Cambridge University Press.
- Morrison M., 1995, „The new aspect: symmetries as meta-laws”, w: F. Weinert (red.), *Laws of nature: essays on the philosophical, scientific and historical dimensions*, Berlin: de Gruyter, s. 157-88.
- Mortimer H., 1982, *Logika indukcji. Wybrane problemy*, Warszawa: PWN.
- Muthen B. O., 1987, „Response to Freedman’s critique of path analysis: improve credibility by better methodological training”, *Journal of Educational Statistics* 12, s. 178-84.
- Mynarski S. (red.), 2002, *Zależności przyczynowo-skutkowe w badaniach rynkowych i marketingowych*, Kraków: WAE.

- Mynarski S., 2003, „Wykorzystanie analizy kowariancji w badaniach zależności przyczynowo-skutkowych”, *Zeszyty Naukowe. Akademia Ekonomiczna w Krakowie* 640, s. 7-20.
- Neapolitan R. E., 2004, *Learning Bayesian networks*, Upper Saddle River: Pearson Prentice Hall.
- Neyman J., 1923/1990, „On the application of probability theory to agricultural experiments”, *Statistical Science* 5, 465-72.
- Neyman J., 1930, *O zasadach metody statystycznej w eugenicie i statystyczne podstawy badań dziedziczności*, Warszawa: PTE.
- Neyman J., 1934, „On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection”, *The Journal of the Royal Statistical Society, Series A* 97, s. 558-606.
- Neyman, J., Iwazskiewicz K., Kołodziejczyk S., 1935, „Statistical problems in agricultural experimentation”, *Supplement to the Journal of the Royal Statistical Society, Series B* 2, s. 107-180.
- Newton I., 1687/1962, *Mathematical principles of natural philosophy*, tłum. A. Motte, Berkeley: University of California Press.
- Niiniluoto I., 1999, *Critical scientific realism*, Oxford: Clarendon.
- Norton J. D., 2003, „Causation as folk science”, *Archives for the Philosophy of Science*, <http://philsci-archive.pitt.edu/archive/00001214>.
- Novick L., Cheng P., 2004, „Assessing interactive causal influence”, *Psychological Review* 111, s. 455-85.
- Nowak L., 1972, *Model ekonomiczny. Studium z metodologii ekonomii politycznej*, Warszawa: PWE.
- Nowak S., 1960, „Some problems of causal interpretation of statistical relationships”, *Philosophy of Science* 27, s. 23-38.
- Nowak S., 1968, „Przyczynowe interpretacje zależności statystycznych w nieeksperymentalnych badaniach socjologicznych”, w: K. Szaniawski (red.), *Metody statystyczne w socjologii. Wybrane zagadnienia*, Warszawa: PWN, s. 91-182.
- Nowak S., 1970, *Metodologia badań socjologicznych*, Warszawa: PWN.
- Nowak S. (red.), 1965, *Metody badań socjologicznych*, Warszawa: PWN.
- Osiewalski J., 2001, *Ekonometria bayesowska w zastosowaniach*, Kraków: AE.
- Palmgren J., 2004, „Introduction to causal modelling and inference”, *Scandinavian Journal of Statistics* 31, s. 159-60.
- Pan W., Frank K., 2003, „A probability index of the robustness of a causal inference”, *Journal of educational and behavioral statistics* 28, s. 315-37.
- Paneth N., 2004, „Assessing the contributions of John Snow to epidemiology. 150 years after removal of the Broad Street pump handle”, *Epidemiology* 15, s. 514-16.

- Paneth N., Vinten-Johansen P., Brody H., Rip M., 1998, „A rivalry of foulness: official and unofficial investigations of the London cholera epidemic of 1854”, *American Journal of Public Health* 88, s. 1545-53.
- Pargetter R., 1984, „Laws and modal realism”, *Philosophical Studies* 46, 335-47.
- Paruolo P., 2005, „Automated inference and the future of econometrics: a comment”, *Econometric Theory* 21, s. 78-84.
- Pawłowski T., 1969, *Metodologiczne zagadnienia humanistyki*, Warszawa: PWN.
- Pawłowski T., 1977, *Pojęcia i metody współczesnej humanistyki*, Wrocław: Ossolineum.
- Pawłowski T., 1978, *Tworzenie pojęć i definiowanie w naukach humanistycznych*, Warszawa: PWN.
- Pearl J., 1988, *Probabilistic reasoning in intelligent systems: networks of plausible inference*, San Mateo: Morgan Kaufmann Publishers.
- Pearl J., 1995a, „Causal diagrams for empirical research”, *Biometrika* 82, s. 669-88.
- Pearl J., 1995b, „Causal inference from indirect experiments”, *Artificial Intelligence in Medicine* 7, s. 561-82.
- Pearl J., 1995c, „Rejoinder to discussions of causal diagrams for empirical research”, *Biometrika* 82, s. 702-10.
- Pearl J., 1998, „Graphs, causality, and structural equation models”, *Sociological Methods and Research* 27, s. 226-84.
- Pearl J., 1999a, „Graphs, structural models, and causality”, w: (Glymour i Cooper), s. 50-138.
- Pearl J., 1999b, „Probabilities of causation: three counterfactual interpretations and their identification”, *Synthese* 121, s. 93-114.
- Pearl J., 2000, *Causality: models, reasoning, and inference*, Cambridge: Cambridge University Press.
- Pearl J., 2002, „Causal inference in the health sciences: a conceptual introduction”, *Health Services and Outcomes Research Methodology* 2, s. 189-220.
- Pearson K., 1892, *The grammar of science*, Londyn: Black.
- Phillips P. C. B., 2005a, „Automated discovery in econometrics”, *Econometric Theory* 21, s. 3-20.
- Phillips P. C. B., 2005b, „Automated inference and the future of econometrics: a colloquium for ET's 20th anniversary”, *Econometric Theory* 21, s. 1-2.
- Phillips P. C. B., 2005c, „HAC estimation by automated regression”, *Econometric Theory* 21, s. 116-42.
- Pietruska-Madej E., Strawiński W. (red.), 1995, *Episteme. Z problemów współczesnej teorii wiedzy*, Warszawa: WFiS.
- Pollard D., 2001, *A user's guide to measure theoretic probability*, Cambridge: Cambridge University Press.
- Popper K. R., 1934/2002, *Logika odkrycia naukowego*, tłum. U. Niklas. Warszawa: PWN.

- Popper K. R., 1962/1992, „Logika nauk społecznych”, w: K. R. Popper, *W poszukiwaniu lepszego świata. Wykłady i rozprawy z trzydziestu lat*, oprac. i tłum. A. Malinowski, Warszawa: KiW, s. 82-101.
- Popper K. R., 1990/1996, *Świat skłonności i dwie nowe koncepcje przyczynowości*, tłum. A. Chmielewski, Kraków: Znak.
- Poterba J. M., 2003, „Some observations on health status and economic status”, *Journal of Econometrics* 112, s. 65-67.
- Pratt J. W., Schlaifer R., 1988, „On the interpretation and observation of laws”, *Journal of Econometrics* 39, s. 23-52.
- Przełęcki M., 1959/1966, „Operacjonizm”, w: T. Pawłowski (red.), *Logiczna teoria nauki*, Warszawa: PWN, s. 99-121.
- Przełęcki M., 1988, *Logika teorii empirycznych*, Warszawa: PWN.
- Psillos S., 1999, *Scientific realism: how science tracks truth*, London: Routledge.
- Psillos S., 2000, „The present state of the scientific realism debate”, *British Journal for the Philosophy of Science* 51, s. 705-28.
- Psillos S., 2002, *Causation and explanation*, Montreal: McGill-Queen's University Press.
- Putnam H., 1965, „Trial and error predicates and the solution to a problem of Mostowski”, *Journal of Symbolic Logic* 30, s. 49-57.
- Putnam H., 1975/1999, „Realism in mathematics and elsewhere”, w: (Klee 1999), s. 317-21. Fragment tekstu „What is a mathematical truth?”.
- Putnam H., 1982, „Three kinds of scientific realism”, *The Philosophical Quarterly* 32, s. 195-200.
- Qin D., Gilbert C. L., 2001, „The error term in the history of time series econometrics”, *Econometric Theory* 17, s. 424-50.
- Quine W. V., 1951/1969, „Dwa dogmaty empiryzmu”, w: W. V. Quine, *Z punktu widzenia logiki. Eseje logiczno-filozoficzne*, tłum. B. Stanosz, Warszawa: PWN, s. 35-70.
- Quine W. V., 1970, „Natural kinds”, w: N. Rescher, A. R. Anderson, P. Benacerraf, A. Grünbaum, G. Massey i R. Rudner (red.), *Essays in honor of Carl G. Hempel. A tribute on the occasion of his sixty-fifth birthday*, Dordrecht: Reidel, s. 5-23.
- Radder H., 2003a, „Technology and theory in experimental science”, w: (Radder 2003c), s. 152-73.
- Radder H., 2003b, „Toward a more developed philosophy of scientific experimentation”, w: (Radder 2003c), s. 1-18.
- Radder H. (red.), 2003c, *The philosophy of scientific experimentation*, Pittsburgh: University of Pittsburgh Press.
- Ramsey F. P., 1925, „Universals”, *Mind* 136, s. 401-17.
- Reichenbach H., 1949, *The theory of probability*, Berkeley: University of California Press.
- Reichenbach H., 1954, *Nomological statements and admissible operations*, Amsterdam: North-Holland.

- Reichenbach H., 1976, *Laws, modalities and counterfactuals*, Berkeley: University of California Press.
- Reichenbach H., 1978, *Selected writings 1909-53*, t. 1-2, Dordrecht: Reidel.
- Report of the Committee for Scientific Inquiries in relation to the cholera epidemic of 1854*, 1855, Londyn: Her Majesty's Stationery Office.
- Richardson B. W., 1936, „John Snow, M.D. A representative of medical science and art of the Victorian area”, w: J. Snow, *Snow on cholera*, Nowy Jork: The Commonwealth Fund, Londyn: Humphrey Milford, Oxford University Press, s. xxiii-xlvi.
- Rindfuss R., Bumpass L., St. John C., 1980, „Education and fertility: implications for the roles women occupy”, *American Sociological Review* 45, s. 431-47.
- Roberts J., 1999, „‘Laws of nature’ as an indexical term: a reinterpretation of Lewis's best-system analysis”, *Philosophy of Science, Proceedings* 66, s. S502-11.
- Roberts J. T., 2004, „There are no laws of the social sciences”. w: *Contemporary debates in philosophy of science*, C. Hitchcock (red.), Oxford: Blackwell Publishing 2004, s. 151-66.
- Roberts J., 2005, „Measurability and physical laws”, *Synthese* 144, s. 433-47.
- Robins J. M., 1987, „A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods”, *Journal of Chronic Diseases* 40, suppl., 139S-61S.
- Robins J. M., 1995, „Discussion of causal diagrams for empirical research by J. Pearl”, *Biometrika* 82, s. 695-98.
- Robins J. M., 1999, „Testing and estimation of direct effects by reparametrizing directed acyclic graphs with structural nested models”, w: (Glymour i Cooper), s. 349-405.
- Robins J. M., 2003, „General methodological considerations”, *Journal of Econometrics* 112, s. 89-106.
- Robins J. M., Greenland S., 1996, „Identification of causal effects using instrumental variables: comment”, *Journal of the American Statistical Association* 91, s. 456-58.
- Robins J. M., Scheines R., Spirtes P., Wasserman L., 2003, „Uniform consistency in causal inference”, *Biometrika* 90, s. 491-515.
- Robins J. M., Wasserman L., 1999a, „On the impossibility of inferring causation from association without background knowledge”, w: (Glymour i Cooper 1999), s. 305-21.
- Robins J. M., Wasserman L., 1999b, „Rejoinder to Glymour and Spirtes”, w: (Glymour i Cooper 1999), s. 333-42.
- Robinson P. M., 2005, „Robust covariance matrix estimation: HAC estimates with long memory/antipersistence correction”, *Econometric Theory* 21, s. 171-80.
- Rogosa D., 1987, „Causal models do not support scientific conclusions: a comment in support of Freedman”, *Journal of Educational Statistics* 12, s. 185-95.
- Rosenbaum P. R., 1987, „The role of a second control group in an observational study”, *Statistical Science* 2, s. 292-316.

- Rosenbaum P. R., 1995, „Discussion of causal diagrams for empirical research by J. Pearl”, *Biometrika* 82, s. 698-99.
- Rosenbaum P. R., 1996, „Identification of causal effects using instrumental variables: comment”, *Journal of the American Statistical Association* 91, s. 465-68.
- Rosenbaum P. R., 2002, *Observational studies*, Nowy Jork: Springer.
- Rosenbaum P. R., Rubin D. B., 1983, „The central role of the propensity score in observational studies for causal effects”, *Biometrika* 70, s. 41-55.
- Rothenberg T. J., 1987, „Comments on Freedman’s paper”, *Journal of Educational Statistics* 12, s. 196-97.
- Rubin D. B., 1974, „Estimating causal effects of treatment in randomized and nonrandomized studies”, *Journal of Educational Studies* 66, s. 688-701.
- Rubin D. B., 1976, „Inference and missing data”, *Biometrika* 63, s. 581-92.
- Rubin D. B., 1977, „Assignment to treatment group on the basis of a covariate”, *Journal of Educational Statistics* 2, s. 1-26.
- Rubin D. B., 1978, „Bayesian inference for causal effects: the role of randomization”, *The Annals of Statistics* 6, s. 34-58.
- Rubin D. B., 1986, „Comment: which ifs have causal answers?”, *Journal of the American Statistical Association* 81, s. 961-62.
- Rubin D. B., 1989, „Bugs, lacunae, and the Minnesota/DC effect: a discussion of H. Wainer’s ‘Eelworms, bullet holes, and Geraldine Ferraro’”, *Journal of Educational Statistics* 14, s. 175-78.
- Rubin D. B., 1990a, „Formal models of statistical inference for causal effects”, *Journal of Statistical Planning and Inference* 25, s. 279-92.
- Rubin D. B., 1990b, „Neyman (1923) and causal inference in experiments and observational studies”, *Statistical Science* 5, s. 472-80.
- Rubin D. B., 1991, „Practical implications of modes of statistical inference for causal effects and the critical role of the assignment mechanism”, *Biometrics* 47, s. 1213-34.
- Rubin D. B., 2001a, „Estimating the causal effects of smoking”, *Statistics in Medicine* 20, s. 1395-1414.
- Rubin D. B., 2001b, „Using propensity scores to help design observational studies: application to the tobacco litigation”, *Health Services and Outcomes Research Methodology* 2, s. 169-88.
- Rubin D. B., 2004a, „Direct and indirect causal effects via potential outcomes”, *Scandinavian Journal of Statistics* 31, s. 161-70.
- Rubin D. B., 2004b, „Reply to discussion”, *Scandinavian Journal of Statistics* 31, s. 196-198.
- Russell B., 1912-13, „On the notion of cause”, *Proceedings of the Aristotelian Society* 13, s. 1-26.
- Sack D. A., Sack B. R., Nair B. G., Siddique A. K., 2004, „Cholera”, *Lancet* 363, s. 223-33.
- Sady W., 2000, *Spór o racjonalność naukową od Poincarégo do Laudana*, Wrocław: FNP.

- Sady W., 2003, „Obiektywna sytuacja generująca model atomu Bohra”, w: W. Krajewski, W. Strawiński (red.), *Odkrycie naukowe i inne zagadnienia współczesnej filozofii nauki*, Warszawa: Semper, s. 105-18.
- Salmon W. C., 1980, „Probabilistic causality”, *The Philosophical Quarterly* 61, s. 50-74.
- Salmon W. C., 1990, *Four decades of scientific explanation*, Minneapolis: University of Minnesota Press.
- Salmon W. C., 1971a, „Explanation and relevance”, w: R. C. Buck, R. S. Cohen (red.), *In Memory of Rudolf Carnap*, Dordrecht: Reidel, s. 27-39.
- Salmon W. C., 1971b, *Statistical explanation and statistical relevance. With contributions by J. G. Greeno and R. C. Jeffrey*, Pittsburgh: University of Pittsburgh Press.
- Salmon W. C., 1994, „Causality without counterfactuals”, *Philosophy of Science* 61, s. 297-312.
- Salmon W.C., 1997, *Causality and explanation*, Nowy Jork: Oxford University Press.
- Salmon W. C., 2001, „Explaining things probabilistically”, *Monist* 84, s. 208-17.
- Savage L. J., 1954, *The foundations of statistics*, Nowy Jork: Wiley.
- Savage L. J., 1962, *The foundations of statistical inference*, Nowy Jork: Wiley.
- Scheffler I., 1967, *The anatomy of inquiry*, Nowy Jork: Alfred A. Knopf.
- Scheffler I., 1982, „Projectibility: a postscript”, *The Journal of Philosophy* 79, s. 334-36.
- Scheines R., 1999, „Estimating latent causal influences”, w: (Glymour i Cooper), s. 425-37.
- Schlick M., 1932/1949, „Causality in everyday life and recent science”, w: H. Feigl, W. Sellars (red.), *Readings in philosophical analysis*, Nowy Jork: Appleton-Century-Crofts, s. 515-39.
- Schulz L., Gopnik A., 2004, „Causal learning across domains”, *Developmental Psychology* 40, s. 162-76.
- Scriven M., 1962, „Explanations, predictions, and laws”, w: *Scientific Explanation, Space, and Time*, H. Feigl, G. Maxwell (red.), *Minnesota Studies in the Philosophy of Science* 3, Minneapolis: University of Minnesota Press, s. 170-230.
- Scriven M., 1975, „Causation as explanation”, *Noûs* 9, s. 3-16.
- Seidenfeld T., 1979, *Philosophical problems of statistical inference*, Dordrecht: Reidel.
- Sen A. K., 1986, Prediction and economic theory, *Proceedings of the Royal Society of London A*, 407, s. 3-23.
- Seneta E., 1987, „Discussion of D. A. Freedman’s ‘As others see us’”, *Journal of Educational Statistics* 12, s. 198-201.
- Shafer G., 1996, *The art of causal conjecture*, Cambridge, MA: MIT Press.
- Shipley B., 1999, „Exploring hypothesis space: examples from organismal biology”, w: (Cooper i Glymour 1999), s. 441-52.
- Siciński A., 1968, Współczynniki zależności zmiennych niemierzalnych, w: K. Szaniawski (red.), *Metody statystyczne w socjologii. Wybrane zagadnienia*, Warszawa: PWN, s. 233-47.

- Simon H. A., 1954, Spurious correlation: a causal interpretation, *Journal of the American Statistical Association* 49, s. 467-92.
- Simon H. A., Iwasaki Y., 1988, „Causal ordering, comparative statistics, and near decomposability”, *Journal of Econometrics* 39, s. 149-73.
- Sims C., 1972, „Money, income and causality”, *American Economic Review* 62, s. 540–552.
- Sims C., 1977, „Exogeneity and causal ordering in macroeconomic models”, w: C. Sims (red.), *New methods in business cycle research*, Minneapolis: Federal Reserve Bank of Minneapolis, s. 23-43.
- Sims C., 1980, Macroeconomics and reality, *Econometrica* 48, 1-48.
- Sloman S., 2005, *Causal models: how people think about the world and its alternatives*, Nowy Jork: Oxford University Press.
- Sloman S., Lagnado D., 2005, „Do we ‘do’?”, *Cognitive Science* 29, s. 5-39.
- Smith R. J., 2005, „Automatic positive semidefinite HAC covariance matrix and gmm estimation”, *Econometric Theory* 21, s. 158-70.
- Smith S., 2002, „Violated laws, ceteris paribus clauses, and capacities”, *Synthese* 130, s. 235-64.
- Smith T. M. F., Sudgen R. A., 1988, „Sampling and assignment mechanisms in experiments, surveys and observational studies”, *International Statistical Review* 56, s. 165-80.
- Smoking and health: report of the advisory committee to the Surgeon General of the Public Health Service*, Waszyngton, D.C.: Public Health Service 1964.
- Snijders T. A. B., Hagenaars J. A., 2001, „Guest editors' introduction to the special issue on causality at work”, *Sociological Methods and Research* 30, s. 3-10.
- Snow J., 1936, *On the mode of communication of cholera*, w: J. Snow, *Snow on cholera*, Nowy Jork: The Commonwealth Fund, Londyn: Humphrey Milford, Oxford University Press, s. 1-138.
- Sobel D., Tenenbaum J., Gopnik A., 2004, „Children’s causal inferences from indirect evidence: backwards blocking and Bayesian reasoning in preschoolers”, *Cognitive Science* 28, s. 303-33.
- Sobel M. E., 1990, „Effect analysis and causation in linear structural equation models”, *Psychometrika* 55, s. 495-515.
- Sobel M. E., 1995, „Causal inference in the social and behavioral sciences”, w: G. Arminger, C. C. Clogg, M. E. Sobel, *Handbook of statistical modeling for the social and behavioural sciences*, Nowy Jork: Plenum Press, s. 1-32.
- Sobel M. E., 1996, „An introduction to causal inference”, *Sociological Methods and Research* 24, s. 353-379.
- Sober E., 1994, „No model, no inference: a bayesian primer on the grue problem”, w: D. Stalker (red.), *Grue! The new riddle of induction*, Illinois: Open Court, s. 225-40.
- Sosa E. (red.), 1980, *Causation and conditionals*, Oxford: Oxford University Press.

- Spirtes P., Glymour C., Scheines R., 1997, „Reply to Humphreys and Freedman’s review of *Causation, prediction, and search*”, *British Journal for the Philosophy of Science* 48, s. 555-68.
- Spirtes P., Glymour C., Scheines R., 2000, *Causation, prediction, and search*, Cambridge, MA: MIT Press.
- Spirtes P., Scheines R., 2005, „Causal inference of ambiguous manipulations”, *Philosophy of Science, Proceedings* 72, s. S833-45.
- Spirtes P., Glymour C., Scheines R., Meek C., Fienberg S., Slate E., 1999, „Prediction and experimental design with graphical causal models”, w: (Glymour i Cooper), s. 65-93.
- Spohn W. (red.), 1991, *Erkenntnis orientated: a centennial volume for Rudolf Carnap and Hans Reichenbach*, Dordrecht: Kluwer Academic Publishers.
- Spohn W., 2002, „Laws, ceteris paribus conditions, and the dynamics of belief”, *Erkenntnis* 57, s. 373-94.
- Stalker D. (red.), 1994, *Grue! The new riddle of induction*, Illinois: Open Court.
- Strawiński W., 2003, Modele przyczynowości według R. Giere’a, w: W. Krajewski, W. Strawiński (red.), *Odkrycie naukowe i inne zagadnienia współczesnej filozofii nauki*, Warszawa: Semper, s. 227-33.
- Styczeń M., 1977, „Two methods of taxonomy”, K. Szaniawski (red.), *Problems of formalization in the social sciences*, Wrocław-Warszawa: Zakład Narodowy im. Ossolińskich-Wydawnictwo PAN, s. 139-62.
- Such J., 1967, „Johna Stuarta Milla koncepcja uniwersalności oraz niezawodności praw”, w: W. Krajewski (red.), *Pojęcie prawa nauki w XIX wieku*, Warszawa: PWN, s. 29-70.
- Such J., 1972, *O uniwersalności praw nauki*, Warszawa: KiW.
- Sulek A., 1979, „Schematy eksperymentalne w badaniach społecznych”, w: A. Sulek (red.), *Logika analizy socjologicznej. Wybór tekstów*, Warszawa: WUW, s. 101-35.
- Suppes P., 1970, *A probabilistic theory of causality*, Amsterdam: North-Holland.
- Suppes P., 1974, *Probabilistic metaphysics*, Uppsala: University of Uppsala Press.
- Sutherland J., 1854, *Report of the General Board of Health on the epidemic cholera of 1848 and 1849*, Londyn: W. Clowes & Sons.
- Swanson N. R., Granger C., 1997, „Impulse response functions based on a causal approach to residual orthogonalization in vector autoregressions”, *Journal of the American Statistical Association* 92, s. 357-67.
- Swinburne R. (red.), 2002, *Bayes’s theorem*, Oxford: Oxford University Press.
- Szaniawski K., 1959/1994, „O indukcji eliminacyjnej”, w: (Szaniawski 1994), s. 148-62.
- Szaniawski K., 1961/1994, „O podstawowych formach wnioskowania statystycznego”, tłum. A. Grobler, w: (Szaniawski 1994), s. 197-207.
- Szaniawski K., 1968, „Indukcja statystyczna w badaniach socjologicznych”, w: K. Szaniawski (red.), *Metody statystyczne w socjologii. Wybrane zagadnienia*, Warszawa: PWN, s. 71-89.

- Szaniawski K., 1994, *O nauce, rozumowaniu i wartościach. Pisma wybrane*, wybór i oprac. J. Woleński, wstęp S. Amsterdamski, J. Woleński, Warszawa: PWN.
- Sztejnberg D., 1931/1990, „O tzw. konieczności związków przyrodzonych”, w: J. Kotarbińska, *Z zagadnień teorii nauki i teorii języka*, Warszawa: PWN, s. 19-58.
- Szymański A., 1996, „Specyfika przenikania matematyki do społecznych badań doświadczalnych”, w: J. Such, M. Szcześniak (red.), *Osobliwości przedmiotowo-metodologiczne w nauce*, Poznań: WFH, s. 123-40.
- Tinbergen J., 1933, „The notions of horizon and expectancy in dynamic economics”, *Econometrica* 1, s. 247-64.
- Tinbergen J., 1939, *Statistical testing of business cycle theories*, Geneva: League of Nations.
- Tinbergen J., 1940, „Econometric business cycle research”, *The Review of Economic Studies* 7, s. 73-90.
- Tintner G., 1938, „A note on economic aspects of the theory of errors in time series”, *The Quarterly Journal of Economics* 53, s. 141-49.
- Tooley M., 1997, *Time, tense and causation*, Oxford: Clarendon.
- Topolski J., 1984, *Metodologia historii*, Warszawa: PWN.
- Toulmin S., 1970, „Reasons and causes”, w: R. Borger, F. Cioffi (red.), *Explanation in the behavioral sciences*, Cambridge: Cambridge University Press, s. 1-26.
- Trout J. D., 2003, *Measuring the intentional world: realism, naturalism, and quantitative methods in the behavioral sciences*, Nowy Jork: Oxford University Press.
- Turner S. P., 1997, „‘Net effects’. A short history”, w: (McKim i Turner 1997), s. 23-45.
- Twardowski K., 1912/1965, „O czynnościach i wytworach”, w: K. Twardowski, *Wybrane pisma filozoficzne*, Warszawa: PWN, s. 217-40.
- van Fraassen B., 1977a, „The pragmatics of explanation”, *American Philosophical Quarterly* 14, s. 143-150.
- van Fraassen B., 1977b, „The only necessity is verbal necessity”, *The Journal of Philosophy* 74, s. 71-85.
- van Fraassen B., 1980, *The scientific image*, Oxford: Clarendon Press.
- van Fraassen B., 1981, „Essences and the laws of nature”, w: R. Healey (red.), *Reduction, time, and reality: studies in the philosophy of the natural sciences*, Cambridge: Cambridge University Press, s. 189-200.
- van Fraassen B., 1984, „Belief and the will”, *The Journal of Philosophy* 81, s. 235-56.
- van Fraassen B., 1989, *Laws and symmetry*, Oxford: Clarendon.
- van Fraassen B., 1993a, „Armstrong, Cartwright, and Earman on laws and symmetry”, *Philosophy and Phenomenological Research* 53, s. 431-44.
- van Fraassen B., 1993b, „Precis of *Laws and symmetry*”, *Philosophy and Phenomenological Research* 53, s. 411-12.
- van Fraassen B., 1994, „The world of empiricism”, w: J. Hilgevoort (red.), *Physics and our view of the world*, Cambridge: Cambridge University Press, s. 114-134.

- van Fraassen B., 1996, „Science, materialism, and false consciousness”, w: J. Kvanvig (red.), *Warrant in contemporary epistemology: essays in honor of Alvin Plantinga's theory of knowledge*, Lanham: Rowman, s. 149-181.
- van Fraassen B., 2002, *The empirical stance*, New Haven: Yale University Press.
- van Fraassen B., 2004a, „Précis of *The empirical stance*”, *Philosophical Studies* 121, s. 127-32.
- van Fraassen B., 2004b, „Replies to discussion on *The empirical stance*”, *Philosophical Studies* 121, s. 171-92.
- van Fraassen B., 2005, „Science as representation: flouting the criteria”, *Philosophy of Science, Proceedings* 72, s. S794-804.
- Verma T., Pearl J., 1990, „Equivalence and synthesis of causal models”, w: P. Boissone, M. Henrion, L. N. Kanal i J. F. Lemmer (red.), *Uncertainty in artificial intelligence*, t. 6, Amsterdam: Elsevier, s. 255-68.
- von Wright G. H., 1971, *Explanation and understanding*, Ithaca: Cornell University Press.
- von Wright G. H., 1974, *Causality and determinism*, Nowy Jork: Columbia University Press.
- von Wright G. H., 1993, „On the logic of the causal relations”, w: *Causation*, E. Sosa, M. Tooley (red.), Oxford: Oxford University Press, s. 105-124.
- Wainer H., 1989, „Eelworms, bullet holes, and Geraldine Ferraro: some problems with statistical adjustment and some solutions”, *Journal of Educational Statistics* 14, s. 121-40.
- Waldmann M., Hagmayer Y., 2005, „Seeing versus doing: two modes of accessing causal knowledge”, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31, s. 216-27.
- Wang J., Roehrl M., 2005, „Anthrax vaccine design: strategies to achieve comprehensive protection against spore, bacillus, and toxin”, *Medical Immunology* 4, s. 1-8.
- Wawrzyniak A., 1964, „Koncepcja stosunku przyczynowego Dawida Hume'a”, *Roczniki Filozoficzne* 12, z. 1, s. 39-51.
- Weber M., 1922, *Gesamelte Aufsätze zur Wissenschaftstheorie*, Tübingen: Mohr-Siebeck.
- Weber M., 1949, *The methodology of the social sciences*, tłum. E. Shils, H. Finch, Londyn: Macmillan.
- Western B., 1999, „Guest editors' introduction to the special issue on Bayesian methods in the social sciences”, *Sociological Methods and Research* 28, s. 3-6.
- Whitehead A. N., 1959, *Symbolism: its meaning and effect*, Nowy Jork: Capricorn Books.
- Whittaker J., 1990, *Graphical models in applied multivariate statistics*, Chichester: Wiley.
- Wierchoń S. T., 1996, *Metody reprezentacji i przetwarzania informacji niepewnej w ramach teorii Dempstera-Shafera*, Warszawa: IPI PAN.
- Williamson J., 1999, „Countable additivity and subjective probability”, *British Journal for the Philosophy of Science* 50, s. 401-16.
- Wiśniewski J., 1992, *Die mathematischen Modelle der Kausalität. Ein Versuch der methodologischen Analyse*, Poznań: WN UAM.

- Wold H., 1938, *A study in the analysis of stationary time series*, Uppsala: Almqvist i Wiksell.
- Wold H., 1949, „Statistical estimation of economic relationships”, *Econometrica* 17, s. 1-22.
- Wold H., 1969, „Econometrics as pioneering in nonexperimental model building”, *Econometrica* 37, s. 369-81.
- Wold H., 1987, „Response to D. A. Freedman”, *Journal of Educational Statistics* 12, s. 202-205.
- Woleński J., 1979/1996, „Kontrowersje metametodologiczne”, w: J. Woleński, *W stronę logiki*, Kraków: Aureus, s. 239-50.
- Woleński J., 1985, *Filozoficzna Szkoła Lwowsko-Warszawska*, Warszawa: PWN.
- Woleński J., 2001, „Naturalizm, antynaturalizm i podstawy statystyki”, *Filozofia Nauki* 33, s. 147-55.
- Woodward J., 1984, „A theory of singular causal explanation”, *Erkenntnis* 21, s. 231-62.
- Woodward J., 1989, „The causal mechanical model of explanation”, *Minnesota Studies in the Philosophy of Science* 13, s. 357-83.
- Woodward J., 1992, „Realism about laws”, *Erkenntnis* 36, s. 181-218.
- Woodward J., 1997, „Explanation, invariance and intervention”, *Philosophy of Science* 64 Supplement, s. S26-S41.
- Woodward J., 1999, „Causal interpretation in systems of equations”, *Synthese* 121, s. 199-247.
- Woodward J., 2000, „Explanation and invariance in the special sciences”, *The British Journal for the Philosophy of Science* 51, s. 197-254.
- Woodward J., 2001, „Law and explanation in biology: invariance is the kind of stability that matters”, *Philosophy of Science* 68, s. 1-20.
- Woodward J., 2002, „Explanation”, w: P. Machamer, M. Silberstein (red.), *The Blackwell guide to the philosophy of science*, Oxford: Blackwell Publishers, s. 37-54.
- Woodward J., 2002c, „What is a mechanism? A counterfactual account”, *Philosophy of Science, Proceedings* 69, s. S366-77.
- Woodward J., 2003a, „Experimentation, causal inference, and instrumental realism”, w: (Radder 2003c), s. 87-118.
- Woodward J., 2003b, *Making things happen: a theory of causal explanation*, Nowy Jork: Oxford University Press.
- Woodward J., 2004, „Counterfactuals and causal explanation”, *International Studies in the Philosophy of Science* 18, s. 41-72.
- Woodward J., Hitchcock C., 2003a, „Explanatory generalizations, part I: a counterfactual account”, *Noûs* 37, s. 1-24.
- Woodward J., Hitchcock C., 2003b, „Explanatory generalizations, part II: plumbing explanatory depth”, *Noûs* 37, s. 181-99.
- Wójcicki R., 1974, *Metodologia formalna nauk empirycznych*, Wrocław-Warszawa: Zakład Narodowy im. Ossolińskich-Wydawnictwo PAN.

- Wójcicki R., 1979, „Models and theories of empirical phenomena”, w: R. Wójcicki, *Topics in the formal methodology of empirical sciences*, Dordrecht-Wrocław: Reidel-Ossolineum, s. 157-163.
- Wójcicki R., 1982, *Wykłady z metodologii nauk*, Warszawa: PWN.
- Wójcicki R., 1991, *Teorie w nauce*, Warszawa: IFiS PAN.
- Wójcicki R., 1995, „Trzy modele transformacji nauki”, w: (Czarnocka 1995), s. 19-44.
- Wright P. G., 1915, „Review of Moore's Economic Cycles. Economic cycles. Their law and cause”, *The Quarterly Journal of Economics* 29, s. 631-41.
- Wright P. G., 1929, „Review of Schultz H., *Statistical laws of demand and supply*”, *Journal of the American Statistical Association* 24, s. 207-15.
- Wright S., 1921, „Correlation and causation”, *Journal of Agricultural Research* 20, s. 557-85.
- Wright S., 1934, „The method of path coefficients”, *The Annals of Mathematical Statistics* 5, s. 161-215.
- Wright S., 1960a, „Path coefficients and path regressions: alternative or complementary concepts?”, *Biometrics* 16, s. 189-202.
- Wright S., 1960b, „The treatment of reciprocal interaction, with or without lag, in path analysis”, *Biometrics* 16, s. 423-45.
- Wynder E., 1954, „Tobacco as a cause of lung cancer”, *The Pennsylvania Medical Journal* 57, s. 1073-83.
- Wynder E., Bross I., Cornfield J., O'Donnell W., 1956, „Lung cancer and women”, *The New England Journal of Medicine* 255, s. 1111-21.
- Wynder E., Graham E., 1950, „Tobacco smoking as a possible etiologic factor in bronchogenic carcinoma”, *The Journal of the American Medical Association* 143, s. 329-38.
- Wynder E., Graham E., 1951, „Etiologic factors in bronchogenic carcinoma with special reference to industrial exposures”, *Archives of Industrial Hygiene and Occupational Medicine* 4, s. 221-35.
- Wynder E., Graham E., Croninger A., 1953, „Experimental production of carcinoma with cigarette tar”, *Cancer Research* 13, s. 855-64.
- Wynder E., Graham E., Croninger A., 1955, „Experimental production of carcinoma with cigarette tar. II. Tests with different mouse strains”, *Cancer Research* 15, s. 445-48.
- Wynder E., Lemon F., Bross I., 1959, „Cancer and coronary artery disease among Seventh-Day Adventists”, *Cancer* 12, s. 1016-28.
- Yule G. U., 1926, „Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series”, *Journal of the Royal Statistical Society* 89, s. 1-63.
- Yule G. U., 1927, „On a method of investigating periodicities in disturbed series, with special reference to Wolfer's sunspot numbers”, *Philosophical Transactions of the Royal Society of London. Series A* 226, s. 267-98.

- Zabell S. L., 1996, „The continuum of inductive methods revisited”, w: J. Earman, J. Norton (red.), *Philosophical problems of the internal and external worlds: essays on the philosophy of Adolf Grünbaum*, Pittsburgh: University of Pittsburgh Press, s. 351-85.
- Zacny B., 2002, „Zagadnienie przyczynowości w modelach ekonometrycznych”, *Ekonometria* 9, s. 57-63.
- Zamecki S., 1967, „O pojęciu prawa w naukach przyrodniczych u Johna Stuarta Milla”, w: W. Krajewski (red.), *Pojęcie prawa nauki w XIX wieku*, Warszawa: PWN, s. 7-28.
- Zeidler P., 1993, *Spór o status poznawczy teorii. W obronie antyrealistycznego wizerunku nauki*, Poznań: WN IF UAM.
- Zellner A., 1971/1996, *An introduction to Bayesian inference in econometrics*, Nowy Jork: Wiley.
- Zellner A., 1988, Causality and causal laws in economics, *Journal of Econometrics* 39, s. 7-21.
- Zieliński A., Mazurek J., 2002, „Badanie związków przyczynowych w epidemiologii”, *Ruch Filozoficzny* 59, s. 275-87.
- Zieliński Z., Bruzda J., 2004, „Porównawcza analiza metod badania odstępów przyczynowych w ekonomii”, w: A. Zeliaś (red.), *Przestrzenno-czasowe modelowanie i prognozowanie zjawisk gospodarczych*, Kraków: AE, s. 9-28.
- Żegleń U., 2000, „Intencjonalność a przyczynowość: problem treści i reprezentacji”, *Kognitywistyka i Media w Edukacji* 1-2, s. 55-72.
- Żegleń U., 2003, *Filozofia umysłu. Dyskusja z naturalistycznymi koncepcjami umysłu*, Toruń: Marszałek.
- Życiński J., 2001, „Kategorie przyczynowości i celowości w filozoficznej interpretacji przyrody”, *Analecta Cracoviensia* 33, s. 283-99.
- Żytkow J. M., 1995, „Automatyzacja odkryć: nowa metodologia i filozofia nauki”, w: E. Pietruska-Madej, W. Strawiński (red.), *Episteme. Z problemów współczesnej teorii wiedzy*, Warszawa: WFiS, s. 43-71.

Skorowidz terminów technicznych¹

- acykliczny graf skierowany – directed acyclic graph
- algorytm odkrycia przyczyn – causal discovery algorithm
 - ~ przeszukiwania – search ~
- analiza czynnikowa – factor analysis
 - konfirmacyjna ~ – confirmatory ~
 - eksplanacyjna ~ – explanatory ~
 - ~ grupowania – cluster ~
 - ~ korelacji cząstkowych – partialling
 - ~ ścieżek – path ~
 - ~ wrażliwości – sensitivity ~
 - ~ wyróżnikowa – discriminant ~
- asercja uzasadniona – warranted assertion
- badania kohortowe – cohort studies
 - ~ obserwacyjne – observational ~
 - ~ typu oddziaływanie-kontrola – case-control ~
- błąd losowy – error term
- cząstkowa/y – partial
 - (ko)wariancja ~ – (co)variance
 - współczynnik korelacji ~ – correlation coefficient
 - współczynnik regresji ~ – regression coefficient

¹ Przyjęta w pracy terminologia uwzględnia rozstrzygnięcia słownikowe i przyjęte w literaturze: w zakresie nauk społecznych – głównie prace pod redakcją naukową J. Brzezińskiego, w zakresie terminologii medycznej i biologicznej – głównie P. Słomskiego (*Słownik lekarski angielsko-polski*, 1991 wyd. 2, Warszawa: PZWL); w zakresie informatyki – monografię P. Cichosza (2000) i prace M. Kłopotka oraz S. Wierzchonia.

- dawka-odpowiedź – dose-response
 diagram przerzedzony – sparse diagram
 dopasowanie – matching
 dostateczność – sufficiency
 dowody empiryczne – empirical evidence
 eksperyment z randomizacją – randomized experiment
 estymator – estimator
 ~ zgodny – consistent ~
 ~ jednorodnie – uniform ~
 ~ punktowo – pointwise ~
 grupa oddziaływania – treatment group
 ~ kontrolna – control ~
 interwencja na (zmienną) X – intervention on (variable) X
 masowa ~ – massive ~
 kolider – collider
 nieosłonięty ~ – unshielded ~
 kondycjonalizacja (warunkowanie w sensie bayesowskim) – conditioning (Bayesian)
 konfunder – confounder
 korelacja reszkowa – residual correlation
 korygowanie (konfundera) – adjustment (for a confounder)
 kryterium tylnych drzwi (Pearl) – back-door criterion
 leczenie zamierzone – intention-to-treat
 manipulacja – manipulation
 mechanizm przypisania – assignment mechanism
 model przyczynowy, Rubina – Rubin causal model
 ~ pajęczynowy – cobweb ~
 ~ potencjalnych rezultatów – potential outcome ~
 ~ autoregresji wektorowej VAR – vector auto-regression ~
 model równań strukturalnych – structural equation model
 możliwość – capacity
 nieprzyczynowość (Granger) – non-causality
 niedoidentyfikowanie parametrów – underidentification of parameters
 niepodatność – non-compliance
 (nie)zależność warunkowa – conditional (in)dependence
 ~ warunkowo względem X – ~ given X
 obciążenie – bias
 ~ selekcją – selection ~
 odpowiedź – response

oddziaływanie – treatment
oszacowanie przyczynowe (Rubin) – causal estimand
plan eksperymentalny – experimental design –
zrandomizowany ~ – randomized ~
planowanie badań – research design
podatność – compliance
pokrycie Markowa – Markov blanket
prawdopodobieństwo brzegowe – marginal probability
prewalencja – prevalence
próba – trial
 medical ~ ~ ~ medyczna
przesianie – screening off
przeszukiwanie punktowe – score based search
 ~ więzowe – constraint based ~
przyczyna aktywna – active cause
 ~ całkowita – total ~
 ~ dopełniająca – contributing ~
 ~ przysposabiająca – predisposing ~
 ~ wspólna – common cause
przyczynowa istotność – causal significance
przyczynowe wnioskowanie – causal inference
przyczynowość międzyjednostkowa – interunit causation
przyczynowy efekt – causal effect
przypadek indeksowy – index case
przypisanie do oddziaływania – treatment assignment
reaktogeniczność – reactivity
realizm co do przedmiotów – entity realism
regresor – regressor
rezultat potencjalny – potential outcome
 ~ zastępczy – surrogate ~
równania równoważne – simultaneous equations
sieć bayesowska – Bayesian Network
skutek, wspólny – common effect
smoła – tar
sprawczość – agency
stabilność – stability (por. założenie o wierności)
ścieżka – path
szlak – trail

- teoria przyczynowości sprawcza – agency theory of causality
termin resztowy – residual term
ukoniecznianie – necessitation
uogólnienie prawopodobne – lawlike generalization
ustawienie wartości X na x – setting the value of X to x
warunkowanie (statyst.) – conditioning (statistical)
wiarogodność – likelihood
wiarygodność – plausibility
wiedza ogólna – background knowledge
wnioskowanie statystyczne – statistical inference
współzmienna – covariate
wynik skłonności – propensity score
~ wyważony – balancing ~
wzorzec – pattern
zachorowalność – incidence
zakłócenie – disturbance
zależność odporna – robust dependence
zależność przyczynowa, pozorna – spurious causal dependence
zależność stabilna – stable dependence
zależność statystyczna – statistical dependence
założenie o stabilności oddziaływania na jednostkę – stable unit treatment value assumption
~ wierności – faithfulness ~
zaniechywalność mocna – strong ignorability
zbiór poprzedników A – ancestral set of A
zdanie kontrfaktyczne – counterfactual
zmienna – variable
~ czysto egzogeniczna – exogenous ~
~ czysto wyjaśniająca – pure explanatory ~
~ dławiąca – suppressor ~
~ oddziaływania – treatment ~
~ odpowiedzi – response ~
~ ukryta – latent ~
~ wskaźnikowa – indicator ~
~ wyjaśniająca – explanatory ~
~ zaobserwowana/obserwowalna – observed/-able ~
~ zniekształcająca – distorter ~
~ zmierzona – measured ~
zmiennosc – variation

Indeks

- Aalen O. O., 196, 403
Abbott A., 118, 125, 403
Achen C. H., 163, 403
Adams P., 71, 403
Adda J., 71, 403
Ajdukiewicz K., 275, 347, 403
Alderson J., 35
Amsterdamski S., 216, 227, 228, 234,
276, 283, 320, 405, 430
Anderson A. R., 424
Anderson D. R., 218, 405
Anderson T. W., 123, 403
Andersson S. A., 384, 403
Angrist J. D., 349, 350, 352-359, 403
Anscombe G. E. M., 18, 261, 404
Arjas E., 118, 196, 404
Arminger G., 428
Armstrong D. M., 79, 246, 247, 248, 249,
261, 404, 430
Arrow K. J., 155, 404
Arystoteles, 20, 21, 276, 278, 280, 404,
409
Ash A., 118, 404
Babington B., 35
Backe A., 319, 322, 404
Barndor-Nielsen O. E., 419
Barnes E., 292, 294, 404
Barnes J., 40
Basmann R. L., 124, 404
Bayes T., 23, 178, 374, 376-378, 381,
404, 408, 409, 412, 429
Beckermann A., 418
Benacerraf P., 424
Bentler P. M., 163, 404
Berger R., 131, 404
Berk R. A., 307, 404
Besnard P., 421
Biela A., 393, 404
Bigaj T., 245, 347, 404, 420
Bing N., 408
Blalock H. M., 154, 405
Bloxam J., 44, 45
Błaszczuk P., 347, 405
Bobyryk J., 261, 405
Bocheński J. M., 275, 276, 405
Bogen J., 14, 405
Bohr N., 66
Boinssone P., 431
Bonk T., 14

- Boreham J., 409
 Borger R., 430
 Bovens L., 14
 Box G., 167, 405
 Boyd R., 332, 405, 410
 Bradshaw G. L., 419
 Braithwaite R., 67, 227, 228, 251, 405
 Brody H., 423
 Bromberger S., 281, 405
 Bronk A., 14, 405, 417
 Bross I., 31, 433
 Brumback B., 148
 Bruzda J., 122, 434
 Brzeziński J., 65, 66, 83, 103, 299, 379, 405, 411, 413, 435
 Buck R. C., 427
 Bujak F., 404
 Bumpass L., 425
 Bunge M., 18, 20, 405
 Burnham K. P., 218, 405

 Campbell D. T., 95, 405, 420
 Carnap R., 217, 280, 284, 331, 406, 427, 429
 Carroll J., 216, 406
 Cartwright N., 14, 18, 257-259, 261, 263, 328, 337, 406, 430
 Chadwick E., 35
 Chambers S. P., 142, 406
 Chandola T., 403
 Cheng P., 80, 406, 422
 Chmielewski A., 424
 Christ C. F., 154, 406
 Chu T., 260, 406
 Churchland P., 406, 411
 Chwedeńczuk B., 419
 Cichosz P., 170, 406, 435
 Cioffi F., 430

 Clements M. P., 169
 Cliff N., 163, 406
 Clogg C. C., 120, 406, 428
 Cochran W. G., 142, 379, 406
 Cohen L. J., 408
 Cohen R. S., 427
 Coleman J. S., 125, 406, 407
 Collingwood R., 260, 338, 407
 Collins J., 226, 407
 Comte A., 23
 Cook S. W., 416
 Cooper G., 407, 412, 423, 425, 427, 429
 Corfield D., 24, 407
 Cox D. R., 66-70, 76, 77, 89, 94, 95, 112, 113, 118-122, 126, 143, 144, 146, 308, 309, 334, 384, 407, 419
 Crámer H., 385, 407
 Croninger A., 31, 433
 Czarnocka M., 407, 433
 Czeżowski T., 20, 24, 228, 231, 262, 276, 277, 319, 407, 416

 D'Agostino R., 139, 407
 Dale I. A., 24, 408
 Danks D., 412
 Darnell A. C., 392, 408
 Davidian M., 139, 420
 Davidson D., 18, 220, 308, 328, 408
 Dawes R., 408
 Dawid A. P., 128, 138, 170, 408
 de Finetti B., 24, 408
 de la Fuente A., 123, 408
 Debreu G., 155
 Dempster A. P., 131, 408
 Deutsch M., 416
 Dobosz M., 123, 408
 Doll R., 31, 80, 408, 409
 Dorling J., 332, 409

- Dowe P., 289, 290, 409
 Dretske F. I., 409
 Duan N., 404
 Duncan D., 154, 409
- Earman J., 24, 228-230, 259-262, 267,
 268, 273, 409, 430, 434
 Edelman E., 18, 409
 Eells E., 79, 409
 Efron B., 131, 409
 Enders W., 167, 409
 Engle R. F., 392, 409
- Farr W., 33, 35, 37, 38, 53, 54, 57, 58, 63,
 65, 66, 72
 Feigl H., 427
 Feldman D., 131, 409
 Fellows P., 128, 410
 Ferguson G. A., 19, 114, 410
 Feyerabend P., 64, 250, 321, 332, 409,
 410
 Fienberg S., 410, 429
 Finch H., 431
 Fisher R. A., 13, 23, 24, 126, 131, 134,
 410
 Fisk P., 364
 Fitelson B., 14
 Fox J., 163, 410
 Frangakis C., 350, 410
 Frank K., 119, 410, 422
 Freedman D. A., 31, 65, 84, 154, 163,
 205, 309, 311, 317, 326, 327, 339,
 392, 393, 399, 400-404, 406, 410, 411,
 415, 419, 421, 425-427, 429, 432
 Friedlander A., 410
 Friedman M., 291, 411
 Frisch R., 163, 411
 Frost W., 32, 33, 34, 54
- Galileusz, 264, 266, 267, 271
 Gatnar E., 122, 186, 411
 Gaul M., 23, 411
 Gawecki B., 18, 411
 Geiger D., 176, 411
 Gelman S., 131, 411
 Ghosh S., 403
 Gibbs P., 410
 Giedymin J., 66, 411
 Giere R., 185, 411, 429
 Gilbert C. L., 147, 164, 424
 Gillies D., 24, 411
 Giza P., 342, 411
 Glaisher J., 35
 Glymour C., 14, 26, 77, 117, 120, 151,
 152, 154, 178, 179, 184, 188, 287,
 307, 326, 327, 339, 379, 398-402, 406,
 407, 410-412, 421, 423, 425, 427, 429
 Goldberger A. S., 153-155, 166, 361-364,
 412
 Goldthorpe J. H., 67-70, 76, 114, 119,
 122-126, 305-312, 334-337, 343, 344,
 412
 Good I. J., 121, 412
 Goodman N., 217-219, 224, 225, 228,
 269, 312, 412, 416
 Gopnik A., 412, 427, 428
 Gordon A. D., 320, 412
 Grabski F., 218, 412
 Graham E., 31, 433
 Granger C., 25, 121-124, 126, 144-146,
 167, 232, 233, 307, 413, 429
 Greenland S., 89, 106-109, 111, 113, 114,
 131, 148, 413, 425
 Greeno J. G., 427
 Grobler A., 327, 339, 413, 429
 Grünbaum A., 424, 434
 Guilford J. P., 120, 413

- Gumański L., 64, 219, 413
 Gutenbaum J., 64, 413
- Haavelmo T., 25, 155-166, 365-370, 408, 413
 Hacking I., 342, 413
 Hagenaars J., 118, 428
 Hagemayer Y., 80, 431
 Hajduk Z., 279, 413
 Hájek A., 14
 Hall B., 35
 Hall N., 407, 413, 422
 Hamilton W. R., 295
 Hankinson R. J., 276, 280, 413
 Hanks S., 421
 Hansen B., 25, 168, 413
 Haritou A., 120, 406
 Harris T., 80, 414
 Hartmann S., 14
 Hassall A., 35, 36, 49, 221
 Hausman D. M., 392, 414
 Healey R., 411, 430
 Heckman J. J., 131, 156, 165-169, 414
 Heller M., 18, 185, 256, 257, 414, 420
 Hempel C. G., 228, 278-286, 291, 292, 294, 295, 299-301, 303, 305, 313, 314, 316, 317, 323, 327, 414, 416, 418, 424
 Hendricks V. F., 112, 414
 Hendry D., 25, 163, 167, 169, 392, 409, 411, 413-415
 Henrion M., 431
 Herbut J., 15
 Hershberger S. L., 147, 415
 Hildreth C., 154, 415
 Hilgevoort J., 430
 Hill B. A., 31, 80, 89-95, 103, 104, 106, 108-115, 309, 346, 408, 409, 415
 Hippocrates, 409
 Hitchcock C., 18, 412, 415, 418, 425, 432
 Hochfeldowa A., 416
 Hoeschele I., 408
 Hofer-Szabó G., 25, 415
 Holland P., 118, 123, 126, 127, 131-138, 145, 146, 154, 307, 308, 336, 415
 Hooker C. A., 406, 411
 Hoover K. D., 25, 415
 Hope K., 163, 415
 Hornowska E., 405, 413
 Horwich P., 112, 415
 Howard-Jones N., 84, 415
 Howarth R. J., 120, 416
 Howie D., 24, 416
 Howson C., 19, 24, 416
 Hume D., 19-23, 89, 251, 294, 416, 431
 Humphreys P., 400-402, 411, 416, 419, 429
 Hurd M., 403
 Hurwicz L., 155, 416
- Imai K., 139, 416
 Imbens G. W., 403, 416
 Ingarden R., 17, 18, 347, 416
 Ivins B., 410
 Iwasaki Y., 309, 334, 428
 Iwaszkiewicz K., 422
- Jadacki J. J., 38, 416
 Jadamus-Hacura M., 167, 416
 Jahoda M., 154, 416
 Jain S., 117, 416
 Janiszowski K., 186, 416
 Jaźwiński J., 218, 412
 Jeffrey R. C., 408, 416, 427
 Jeffreys H., 23
 Jenkins G., 167, 405
 Jensen F. V., 170, 416

- Jeon Y., 167, 413
Jones M., 406
Judycki S., 19, 219, 332, 416, 417
Juhl C., 117, 332, 417
- Kaluszyńska E., 276, 332, 417
Kamiński S., 21, 23, 63, 275, 326, 405, 417
Kamlah A., 25, 417
Kanal L. N., 431
Kant I., 19
Kaprio J., 80, 417
Karhausen L. R., 24, 417
Karlin S., 163, 417
Karpiński J., 18, 76, 147, 342, 417
Kaufman J. S., 417
Kawalec A., 15
Kawalec P., 15, 19, 23, 25, 27, 184, 217, 280, 285, 381, 417, 418
Kelly K., 332, 412, 417, 418
Kelvin W. T., 242
Keynes J., 23
Kiczuk S., 20, 418
Kiiveri H., 25, 384, 418
Kim J., 215, 418
Kincaid H., 418
Kistler M., 234, 418
Kitcher P., 286, 291-296, 313, 418
Klee R., 418, 424
Klein L. R., 155, 166, 418
Klüppelberg C., 419
Kłopotek M. A., 170, 190, 418, 435
Kmita J., 215, 277, 281, 321, 418
Kneale W., 245, 418
Koch R., 32, 92, 221, 271, 317, 409, 415
Kołodziejczyk S., 422
Koopmans T., 25, 155, 163, 364, 418
Korb K., 401, 402, 419
- Koskenvuo M., 80, 417
Kotarbińska J., 430
Kotarbiński T., 419
Kourany J., 14
Krajewski W., 18, 19, 64, 195, 216, 251, 262, 275, 278, 289, 297, 419, 427, 429, 434
Kripke S., 245, 419
Krolzig H. M., 25, 414
Kuersteiner G., 25, 419
Kuhn T., 220
Kushnir T., 412
Kvanvig J., 431
Kwiatkowski T., 21, 419
- Lad F., 176, 371, 419
Lagnado D., 419, 428
Lagrange J. L., 295
Lalonde R., 350, 419
Langley P., 25, 419
Lauritzen S., 26, 170, 174-177, 179-184, 193, 196, 205, 209, 384, 419
Lazarsfeld P. F., 101, 123, 419
Lea J., 59
Leamer E., 392, 419
Leeb H., 25, 419
Lemmer J. F., 431
Lemon F., 31, 433
Lennox J., 14
Leplin J., 419
Levi I., 420
Lewis D., 76, 226, 230, 231, 233, 239, 242, 298, 346, 413, 419, 420, 425
Lewis G., 169
Lieberson S., 308, 311, 312, 393, 420
Liebig J., 37
Lindley D. V., 420
Linscott M., 410

- Little D., 420
 Little R., 350, 420
 Liu T. C., 166, 167, 420
 Losee J., 21, 276, 420
 Lubański M., 23, 420
 Lunceford J. K., 139, 420

 Łukasiewicz J., 245, 416, 420

 Mach E., 23
 Machamer P., 432
 Machowski A., 23, 411
 Mackie J. L., 19, 22, 261, 308, 420
 Maddala G. S., 392, 420
 Madigan D., 403
 Majdański S., 275, 420
 Malewski A., 276, 420
 Malinowski A., 424
 Marciszewski W., 189, 275, 420
 Markowitz H., 155
 Marmot M., 403
 Marshall A., 149, 257, 420
 Massey G. J., 424
 Maxwell G., 427
 Maxwell J. C., 291
 Mazierski S., 216, 245, 420
 Mazurek J., 89, 90, 91, 93, 317, 346, 434
 Mączka J., 76, 420, 421
 McFadden D., 403
 McKim V. R., 406, 410, 421, 430
 Mealli F., 350, 421
 Meek C., 176, 421, 429
 Meggle G., 413
 Mejbaum W., 23, 421
 Mellor D. H., 328, 421
 Mendes P., 408
 Menzies P., 19, 254, 260-263, 339, 421
 Merrill A., 403

 Metallmann J., 18, 20, 75-77, 275, 281, 347, 420, 421
 Meyerson É., 75, 347
 Michalik A., 414
 Mill J. S., 89, 239, 421, 429, 434
 Mizon G., 167, 414
 Mochon D., 147, 421
 Modigliani F., 155
 Mokrzycki E., 20, 279, 421
 Moore H. L., 433
 Morabia A., 65, 421
 Morgan M., 163, 411, 415, 421
 Morrison M., 259, 421
 Mortimer H., 25, 112, 127, 421
 Mostowski A., 424
 Motte A., 422
 Muthen B. O., 163, 421
 Mynarski S., 120, 195, 411, 421, 422

 Nair B. G., 426
 Neapolitan R. E., 422
 Newton I., 21, 22, 227, 257, 267, 283, 291, 294, 295, 422
 Neyman J., 23, 118, 126-132, 156, 162, 322, 346, 422, 426
 Niiniluoto I., 161, 332, 333, 422
 Nimitz C., 418
 Normand S.-L. T., 404
 Norton J., 422, 434
 Novick L., 80, 422
 Nowak L., 147, 215, 277, 321, 418, 422
 Nowak S., 66, 281, 299, 303, 416, 419, 422

 O'Donnell W., 31, 433
 Oppenheim P., 278, 414
 Osherson D., 416
 Osiewalski J., 24, 422

- Palmgren J., 196, 422
Pan W., 119, 422
Paneth N., 32, 34-38, 44, 48, 54, 73, 85,
317, 329, 422, 423
Pargetter R., 234, 237, 245, 423
Parner J., 196, 404
Paruolo P., 25, 423
Pasteur L., 37
Paul L., 131, 407
Pawlowski T., 87, 419, 423, 424
Pearl J., 25-27, 118, 125, 147, 151, 152,
159, 161, 170, 175, 176, 178-181, 183,
187-195, 287, 310, 325, 326, 328, 365,
379, 380, 381, 383-393, 407, 408, 410,
411, 416, 423, 425, 426, 431
Pearson E., 156, 162
Pearson K., 23, 408, 422, 423
Perlman M., 403
Petitti D. B., 326, 411
Peto R., 409
Phillips P. C., 25, 423
Pietruska-Madej E., 423, 434
Pitt M., 410
Plantinga A., 431
Platon, 276
Pollard D., 423
Poole C., 131, 413, 417
Popper K. R., 18, 279, 280, 423
Poterba J. M., 71, 424
Pötscher B. M., 25, 419
Pratt J. W., 131, 424
Price H., 13, 14, 19, 254, 260-263, 339,
342, 421
Przełęcki M., 225, 277, 332, 424
Psillos S., 235, 237, 248, 257-259, 279,
280, 283, 286, 289, 290, 328, 332, 424
Putnam H., 25, 332, 424
Qin D., 147, 164, 424
Quine W. V., 219, 220, 424
Radder H., 424, 432
Ramsey F. P., 24, 231, 239, 424
Rédei M., 415
Reichenbach H., 24, 25, 228, 331, 406,
415, 417, 424, 425, 429
Rescher N., 416, 424
Ribeiro T., 403
Richard J. F., 409
Richardson B. W., 31, 34, 314, 425
Richardson T., 412
Richardson T. S., 403
Rindfuss R., 398, 399, 401, 425
Rip M., 423
Roberts J., 260-262, 409, 425
Robins J. M., 131, 179, 192, 393, 425
Robinson P. M., 119, 425
Roehrl M., 205, 431
Rogosa D., 163, 425
Rosenbaum P. R., 96, 121, 131, 138-143,
306, 353, 425, 426
Rossi C., 410
Rothenberg T. J., 163, 426
Rothman K., 89, 106-109, 111, 113, 114
Royer J., 416
Rubin D. B., 11, 117, 118, 126-134, 138-
143, 146, 165, 196, 198-207, 209, 210,
212, 227, 322, 327, 350, 352, 353,
403, 407, 410, 416, 421, 426, 436
Rudner R. S., 424
Russell B., 23, 256, 281, 338, 426
Sack B. R., 426
Sack D. A., 83, 221, 314, 426
Sady W., 66, 426, 427

- Salmon W. C., 277, 278, 283, 285-290,
314, 318, 325, 327, 329, 416, 418, 427
- Savage L. J., 427
- Scheffler I., 427
- Scheines R., 14, 117, 177-179, 187, 188,
287, 326, 379, 398-402, 406, 410, 412,
425, 427, 429
- Schlaifer R., 131, 424
- Schlick M., 23, 228, 427
- Schultz H., 361, 362-364, 433
- Schulz L., 412, 427
- Scriven M., 282, 283, 427
- Seidenfeld T., 24, 427
- Sen A. K., 124, 427
- Seneta E., 163, 427
- Shafer G., 176, 427
- Shils E., 431
- Shiple B., 393-398, 427
- Siciński A., 123, 427
- Siddique A. K., 426
- Silberstein M., 432
- Simon H. A., 35, 126, 155, 309, 334, 419,
428
- Simon J., 35
- Sims C., 167, 232, 428
- Slate E., 429
- Slovan S., 147, 419, 421, 428
- Słomski P., 435
- Słupecki J., 420
- Smith S., 131, 409, 428
- Snijders T. A. B., 118, 428
- Snow J., 15, 26, 31, 32, 34-36, 38-61, 63-
67, 70-76, 79, 80, 82-86, 90, 92, 104,
112, 220-224, 262, 271, 314-317, 323-
325, 329, 333, 337, 340, 341, 425, 428
- Sobel D., 412, 428
- Sobel M. E., 126, 131, 308, 428
- Sober E., 218, 428
- Sosa E., 428, 431
- Sowden L., 420
- Speed T., 25, 129, 418
- Spirtes P., 14, 77, 84, 117, 148, 171, 177-
179, 181, 185-188, 287, 326, 385, 398-
402, 406, 410, 412, 425, 429
- Spohn W., 406, 412, 417, 429
- St. John C., 425
- Stadler F., 415
- Stalker D., 428, 429
- Stanley J. C., 95, 405
- Stanosz B., 424
- Strawiński W., 185, 423, 427, 429, 434
- Styczeń M., 296, 429
- Such J., 228, 238, 429, 430
- Sudgen R. A., 131, 428
- Sułek A., 18, 80, 82, 429
- Suppes P., 121, 429
- Susser M., 31, 65, 89, 94-96, 98-106, 113,
317
- Sutherland I., 33, 37, 38, 409, 429
- Swanson N. R., 232, 233, 429
- Swinburne R., 24, 381, 429
- Szabó L., 415
- Szaniawski K., 24, 422, 427, 429, 430
- Szcześniak M., 430
- Szteinbarg D., 234, 347, 430
- Szubka T., 417
- Szymański A., 299, 430
- Ślaga S. W., 420
- Takane Y., 19, 114, 410
- Tenenbaum J., 428
- Thomson R., 35, 36
- Tinbergen J., 430
- Tintner G., 147, 430
- Tobin J., 155

- Tooley M., 246, 248, 251, 430, 431
Topolski J., 276, 280, 420, 430
Toulmin S., 308, 430
Trout J. D., 430
Turner S. P., 317, 406, 410, 421, 430
Twardowski K., 346, 416, 430
- Urbach P., 19, 24, 416
- van Dyk D. A., 139, 416
van Fraassen B., 24, 233, 234, 237, 240, 242, 246, 247, 248, 249, 259, 260, 263-273, 286, 322, 331, 333, 336-338, 343, 345, 406, 409, 418, 430, 431
Verma T., 176, 187, 385, 411, 431
Vichi M., 320, 412
Vinten-Johansen P., 423
von Wright G. H., 27, 261, 262, 275, 282, 342, 343, 431
- Wainer H., 379, 426, 431
Waldmann M., 80, 431
Walesa C., 404
Wallace C., 401, 402, 419
Wang J., 205, 431
Wasserman L., 425
Wawrzyniak A., 22, 431
Weber M., 277, 321, 431
Weinert F., 421
Wermuth N., 68, 70, 76, 77, 89, 94, 95, 112, 113, 118-122, 126, 143, 144, 146, 308, 334, 384, 407
Wessles U., 413
Western B., 118, 431
Weyl H., 264, 272
Whitehead A. N., 337, 431
Whittaker J., 170, 384, 431
Wicksell K., 366-370
- Wierchoń S. T., 170, 190, 418, 431, 435
Williamson J., 14, 18, 24, 407, 431
Wiśniewski J., 431
Wold H., 163, 364, 432
Woleński J., 14, 24, 281, 347, 430, 432
Woodward J., 82, 251-257, 277, 283, 290, 295-305, 313, 318, 325, 327-329, 392, 414, 432
Working H., 364
Worrall J., 14
Wójcicki R., 23, 25, 64, 215, 216, 332, 432, 433
Wright P. G., 362, 363, 433
Wright S., 23, 148-154, 166, 187, 342, 361-364, 383, 433
Wynder E., 31, 80, 433
- Yule G. U., 433
- Zabell S. L., 23, 434
Zacny B., 122, 434
Zagrodzki M., 410
Zakrzewska M., 405, 413
Zamecki S., 239, 434
Zamiara K., 410
Zeidler P., 434
Zeisel H., 84, 205, 411
Zeliaś A., 434
Zellner A., 124, 434
Zieliński A., 89, 90, 91, 93, 317, 346, 434
Zieliński Z., 122, 434
Ziemińska R., 417
Ziemiński I., 417
Znamierowski C., 416, 421
- Żegleń U., 18, 19, 219, 220, 261, 417, 434
Życiński J., 185, 275, 414, 434
Żytkow J., 26, 419, 434