

Spis treści

WSTĘP	9
CZĘŚĆ I	
ŹRÓDŁA ETYKI R. M. HARE’A	
§ 1 ŻYCIE I DOROBIEK R. M. HARE’A	17
§ 2 INSPIRACJE FILOZOFICZNE	22
§ 2. 1 G. E. Moore – błąd naturalistyczny	22
§ 2. 2 Emotywizm	26
§ 2. 3 Szkoła analizy języka potocznego	31
§ 2. 4 Kant i utilitaryzm	36
CZĘŚĆ II	
METAETYKA	
§ 3 NONDESKRYPTYWIZM R. M. HARE’A	44
§ 3. 1 Geneza i funkcja języka moralnego	45
§ 3. 2 Klasyfikacja teorii etycznych	46
§ 3. 3 Krytyka teorii konkurencyjnych	55
§ 3. 3. 1 Krytyka emotywizmu	55
§ 3. 3. 2 Reductio ad absurdum deskryptywizmu	57
§ 3. 3. 3 Krytyka teorii realistycznych	65
§ 3. 3. 4 Krytyka scjentyzmu – Hare’a zmodyfikowany test otwartego pytania	69
§ 3. 4 Kryteria oceny teorii metaetycznych	72
§ 4 PRESKRYPTYWIZM	79
§ 4. 1 Preskrypcje	79
§ 4. 1. 1 Analiza gramatyczna sądów moralnych	81

§ 4. 1. 2 Analiza w języku teorii czynności mowy	83
§ 4. 2 Czynności mowy a znaczenie	86
§ 4. 3 Zarzut błędu redukcji sądów moralnych do zaleceń	88
§ 4. 4 Zarzut błędu czynności mowy	89
§ 4. 5 Odpowiedź Hare'a na argument błędu czynności mowy	94
§ 4. 6 Znaczenie a kryterium użycia	96
§ 4. 7 Argument Fregego – Geacha	100
§ 4. 8 Hare'a teoria czynności mowy	103
§ 4. 9 Konwencje a intencje	106
§ 4. 10 Logika moralnej gry językowej a znaczenie	107
§ 4. 11 Logika imperatywów	111
§ 4. 12 Problem przejścia od „jest” do „powinien”	121
§ 4. 13 Internalizm R. M. Hare'a	129
§ 4. 13. 1 Słabość woli	134
§ 4. 13. 2 Sceptyczny amoralista	144
§ 4. 14 Nadrzędność	148
§ 5 UNIWERSALNY PRESKRYPTYWIZM	152
§ 5. 1 Uzasadnienie	152
§ 5. 2 Uniwersalność	154
§ 5. 3 Prawda i obiektywność	163

CZĘŚĆ III

ETYKA NORMATYWNA

§ 6 ARCHANIOŁ – WZÓR MYŚLENIA MORALNEGO	175
§ 6. 1 Dwie płaszczyzny myślenia moralnego	175
§ 6. 2 Od metaetyki do myślenia krytycznego	178
§ 6. 3 Myślenie krytyczne: teoria cech istotnych i proces ich uniwersalizacji	182
§ 6. 3. 1 Pierwszy etap myślenia krytycznego	184
§ 6. 3. 2 Drugi etap myślenia krytycznego	186
§ 6. 3. 2. 1 Fanatyk	190
§ 6. 3. 3 Trzeci etap myślenia krytycznego	191
§ 6. 3. 4 Racjonalność	197
§ 6. 4 Myślenie krytyczne a tożsamość osobowa	212
§ 6. 5 Utylitaryzm myślenia krytycznego	214

§ 7 CZŁOWIEK A MYŚLENIE KRYTYCZNE	219
§ 7. 1 Czy utylitaryzm jest zbyt wymagający?	224
§ 7. 1. 1 Czyn zły a czyn niesłuszny	225
§ 7. 1. 2 Obowiązki supererogacyjne i integralność ludzkiego życia	227
§ 7. 1. 3 Utylitaryzm a stronniczość	231
§ 7. 2 Utylitaryzm a sprawiedliwość	232
§ 7. 3 Stabilność dwóch poziomów myślenia moralnego	235
§ 7. 4 Cnoty a utylitaryzm	236
§ 7. 5 Utylitaryzm a myślenie intuicyjne	238

ZAKOŃCZENIE

Bibliografia	256
--------------	-----

Wstęp

Richard Mervyn Hare zajmuje we współczesnej filozofii moralności pozycję szczególną. Uważany jest za jednego z najwybitniejszych etyków ostatniego stulecia. Już pierwsza jego książka¹, wydana w 1952 roku, przyczyniła się do nakreślenia ram dyskusji metaetycznych w epoce powojennej. Richard Brandt w krytycznym eseju poświęconym dorobkowi Hare'a² stwierdził: „Pisarstwo etyczne profesora Hare'a [...] stanowi najbardziej solidne i wyrafinowane przedstawienie i obronę utilitaryzmu od czasów Henry'ego Sidgwicka”³. Uczeń Hare'a, Peter Singer, poszedł jeszcze dalej: „kiedy przyjdzie czas na napisanie historii dwudziestowiecznej etyki, wierzę, że właśnie praca Hare'a będzie postrzegana jako wnosząca największy do niej wkład.”⁴ Wydaje się, iż w znacznej mierze opinie te są zasłużone. Wystarczy przyjrzeć się akademickiej karierze Hare'a, jego dorobkowi oraz liczbie odwołań do jego prac. Jednakowoż śledząc toczącą się od dziesięciu lat dyskusję metaetyczną w krajach anglosaskich, zauważyć można, iż zainteresowanie jego teorią znacznie spadło. Choć powstały koncepcje metaetyczne, które są spokrewnione z Hare'a uniwersalnym preskrytywizmem (projektywizm S. Blackburna⁵, norm-ekspresy-

¹ R. M. Hare, *The Language of Morals*, Oxford University Press, London 1964.

² Tenże, *Myslenie moralne. Jego płaszczyzny, metody i istota*, tłum. J. Margański, Fundacja Aletheia, Warszawa 2001.

³ R. B. Brandt, *Act-Utilitarianism and Metaethics*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics. Essays on „Moral Thinking”*, Clarendon Press, Oxford 1988, s. 27.

⁴ P. Singer, *R. M. Hare's Achievements in Moral Philosophy. A talk for the Memorial Service at St Mary's Church, Oxford, 25 May, 2002*, „Utilitas” (3) 2002, s. 309.

⁵ Simon Blackburn jest twórcą tzw. **projektywizmu**. Zob. S. Blackburn, *Spreading the Word*, Oxford University Press, Oxford 1984. Twierdzi on, iż jeśli nie istnieją emotywne predykaty, które jawią się jako obiektywne, to powinniśmy je

wizm A. Gibbarda⁶ czy teoria kognitywistycznego ekspresywnizmu M. Timmonsa⁷), na plan pierwszy wysuwają się koncepcje, których był on nieustępliwym krytykiem (np. naturalizm). Przeglądając popularny, wydany w 2003 roku, wstęp do metaetyki Alexandra Millera *An Introduction to Contemporary Metaethics*⁸, ze zdziwieniem można zauważyć, iż nie uwzględnia się w nim dorobku Hare'a.

Richard Hare dziedzinę rozważań związanych z moralnością podzielił na etykę, etykę opisową i moralność⁹. Znaczenia, jakie im nadawał (z wyjątkiem etyki opisowej), nie odpowiadają jednak najczęściej spotykanym ich definicjom w literaturze przedmiotu. Przez etykę Hare rozumiał bowiem teorię znaczenia wypowiedzi etycznych i z tego powodu to, co określił tym mianem, nazwiemy metaetyką, a to, co nazwał moralnością – etyką normatywną. Do

wynaleźć. Poszukiwania zacząć musimy nie od terminów moralnych takich jak „dobry” czy „zły”, lecz od wykrzyknień – języka emotywnego. Tylko za sprawą takiego języka możemy bowiem wyrażać nasze postawy. Potrzebujemy „instrumentu do poważnej, podlegającej namysłowi (*reflective*), oceniającej praktyki, pozwalającej nam na borykanie się z poprawianiem postaw, ich zderzaniem się, implikacjami i koherencją”; tamże, s. 195. Tak więc musimy „wynaleźć predykaty odpowiadające postawom i traktować *zobowiązania* (*commitments*), tak jak gdyby były one ocenami, a następnie użyć wszystkich naturalnych sposobów, aby wynegocjować prawdziwość owych sądów”; tamże. Zdaniem Blackburna wartości moralne są więc quasi-zobiektywizowanymi uczuciami; tamże, s. 182-189.

⁶ Alan Gibbard jest autorem tzw. **norm-ekspresywnizmu** (*norm-expressivism*). Zob. A. Gibbard, *Wise Choices, Apt Feelings*, Clarendon Press, Oxford 1990. Jego zdaniem sądy normatywne służą do koordynacji działań i uczuć. Zasady moralne rozstrzygają, które z moralnych uczuć są racjonalne czy uzasadnione. Nazywać, generalnie rzecz ujmując, czyjeś uczucie uczuciem o dobrych podstawach (*warranted*), to wyrazić akceptację normy, która na nie zezwala; tamże, s. 154. Normy rozumiane są jako nakazy lub imperatywy, tamże, s. 70. Wyznaczają one racjonalność uczuć moralnych, takich jak: poczucie winy, gniew czy resentment; tamże, s. 52. Zdaniem Gibbarda oceny muszą być formułowane w języku uczuć, gdyż tylko wtedy są dla nas zrozumiałe; tamże, s. 131. Moralność to zbiór moralnych uczuć, które z perspektywy pełnego i bezstronnego zaangażowania racjonalnie jest posiadać; tamże, s. 128. Sąd normatywny jest ekspresją stanu umysłu, nie jest w ścisłym sensie przekonaniem. Nie jest on także uczuciem czy dyspozycją do uczuć. Jest to raczej złożony stan będący akceptacją odpowiednich norm, które regulują naszymi przekonaniem, działaniami czy uczuciami.

⁷ Zob. M. Timmons, *Morality Without Foundations. A Defense of Ethical Contextualism*, Oxford University Press, New York 1999.

⁸ A. Miller, *An Introduction to Contemporary Metaethics*, Polity Press, Oxford 2003.

⁹ Zob. s. 39-40.

powyższego podziału dodać można również etykę stosowaną. Przez **etykę normatywną** będziemy rozumieć zbiór ocen moralnych wraz z kryteriami ich wyznaczania¹⁰. Jest ona „namysłem nad sposobem sformułowania i obrony prawdziwości zupełnego i ekonomicznego zbioru ogólnych zasad etycznych, a także pewnych zasad mniej ogólnych, które są istotne dla «zbudowania etycznych podstaw» ważniejszych ludzkich instytucji”¹¹. Do przykładowych pytań normatywnych należą: „Czy aborcja jest moralnie zła?”, „Czy powinniśmy zostać wegetarianami?”, „Czy istnieją wojny sprawiedliwe?”, „Jakie cnoty powinniśmy kultywować?”. W skład etyki normatywnej, oprócz odpowiedzi na powyższe pytania, wchodzi również ich uzasadnienie. Przykładem teorii normatywnych jest utylitaryzm czynu, utylitaryzm reguł, utylitaryzm Milla, hedonizm Epikura, etyka kantowska, etyka tomistyczna¹². Ich badanie jest celem i dziedziną rozważań metaetycznych. Przez **metaetykę** rozumieć będziemy dziedzinę filozofii, zajmującą się logiką moralności, stawiającą pytania dotyczące znaczenia sądów i pojęć moralnych (np. „powinno się”, „dobry”, „słuszny”), bytów, do których te pojęcia się odnoszą, sposobu ich poznania. Zadaniem metaetyki jest również wyznaczanie kryteriów oceniania teorii normatywnych oraz metaetycznych. Wśród związanych z nią pytań znaleźć można zagadnienia odnoszące się do¹³: języka (Jakie jest znaczenie sądów moralnych? Czy ich funkcją jest wyrażanie stwierdzeń, czy emocji?); metafizyki (Czy istnieją własności i fakty moralne? Jaki mają charakter?); epistemologii (Czy istnieje wiedza moralna? W jaki sposób możemy się dowiedzieć, które sądy moralne są prawdziwe?); fenomenologii (W jaki sposób są doświadczane własności moralne?); psychologii (Jaka jest relacja między motywacją a obowiązkiem? Czym jest słabość woli?); obiektywności (Czy sądy moralne mogą być prawdziwe/fałszywe?). Tezą metaetyczną jest np.: „Sądy moralne są uniwersalnymi i nadrzędnymi preskrypcjami”, etyczną zaś: „Eutanazja jest niesłuszna i jeśli chcesz to uzasadnić, posłuż się taką a taką wersją utylitaryzmu reguł”. Częścią etyki normatywnej jest etyka stosowana. Mianem *etyki stosowanej* nazywać

¹⁰ Zob. tamże, s. 40.

¹¹ R. B. Brandt, *Etyka. Zagadnienia etyki normatywnej i metaetyki*, tłum. B. Stanosz, Wydawnictwo Naukowe PWN, Warszawa 1996, s. 19.

¹² Por. A. Miller, *An Introduction...*, dz. cyt., s. 1.

¹³ Por. tamże, s. 2.

będziemy taki zbiór pism i problemów, w których wykorzystując dwie pierwsze dziedziny filozofii moralnej, próbuje się odpowiedzieć na pytania dotyczące wartości konkretnych rodzajów czynów. W dziedzinie tej Hare uważany jest za pioniera.¹⁴ Już w 1955 roku napisał pierwszą pracę będącą próbą rozwiązania konkretnego problemu etycznego, tj. problemu posłuszeństwa prawu. W późniejszych latach zaprezentował swoje stanowisko na temat większości praktycznych problemów moralnych: kwestii związanych z prawem, polityką, edukacją, religią, medycyną i nauką.

Głównym celem niniejszej książki jest przedstawienie, poddanie analizie oraz krytyczne ocenie poszczególnych, najważniejszych elementów filozofii moralności Hare'a. Wyznaczenie Hare'owi odpowiedniego miejsca pośród teorii zarówno metaetycznych jak i etycznych, kwestie metafizyczne, będą w niej miały przewagę nad normatywnymi zagadnieniami etycznymi. Najważniejsze pytania, wyznaczające drogę, którą będzie przebywać nasza myśl, mają następującą treść: Czy współcześnie, biorąc pod uwagę liczne, zaprezentowane poniżej argumenty krytyczne skierowane w stronę uniwersalnego preskrytywizmu i utilitaryzmu, warto bronić tych teorii? Jaka jest relacja między rozważaniami metaetycznymi a normatywną metodą myślenia moralnego zaproponowaną przez Hare'a? Czy Richard Hare próbował ufundować etykę normatywną na emotywistycznych założeniach metaetycznych, a jeśli tak, to z jakim skutkiem?

Odpowiedzi na powyższe pytania zostaną udzielone w trzech częściach, podzielonych na rozdziały. W pierwszej części umieścimy koncepcję Hare'a w kontekście historycznym, uwypuklając źródła i założenia leżące u podstaw jego myślenia. Otworzy ją §1, w którym opiszemy życie i dorobek Hare'a. W §2 przedstawimy filozofów, nurty i argumenty, które wpłynęły na jego sposób myślenia: główne problemy metaetyczne zarysowane przez G. E. Moore'a, współczesne Hare'owi teorie emotywistyczne, metodę filozofowania reprezentowaną przez oksfordzką szkołę analizy języka potocznego, której był on przedstawicielem. Porównamy również zwiąże jego etykę z największymi filozofami, których myśli przenikają jego preskrytywizm. Część II poświęcona jest zagadnieniom metaetycznym. Rozpocznie ją §1, którego celem jest zapoznanie czytelnika z różnymi odmianami teorii metaetycznych, uwydatnienie ich sła-

¹⁴ Zob. P. Singer, *R. M. Hare's Achievements...*, dz. cyt., s. 316.

bości oraz wyznaczenie zadań i ewentualnych problemów leżących przed obrońcą uniwersalnego preskrytywizmu. Stanowić on będzie szczegółową analizę tych rozwiązań Hare'a, dzięki którym starał się on sprostać wymogom stawianym dobrym teoriom metaetycznym. W §4 zanalizujemy cechy i problemy związane z preskrytywnością sądów moralnych, w §5 – z ich uniwersalizowalnością. Następnie, w części III, rozpoczynając od opisu struktury jego teorii, wprowadzimy czytelnika w zagadnienia relacji pomiędzy metaetyką a etyką normatywną. W §6 przedstawimy idealną metodę rozwiązywania problemów moralnych, z położeniem nacisku na sposób jej ufundowania na podstawie tez metaetycznych. §7 poświęcony będzie natomiast tym aspektom utilitaryzmu Hare'a, które czynią go metodą zalecaną w codziennym namyśle moralnym. W szczególności sposób zaakcentujemy dwa poziomy myślenia moralnego oraz przedstawimy klasyczne zarzuty stawiane teoriom utilitarystycznym i Hare'a sposób udzielenia na nie odpowiedzi.

W języku polskim nie istnieje kompletne omówienie powyższych zagadnień, a literatura dotycząca zarówno uniwersalnego preskrytywizmu, jak i utilitaryzmu jest raczej uboga. Pewne, niewielkie fragmenty §3.2 i §4.4 oraz całość §4.13 jako artykuły ukazały się drukiem przed publikacją niniejszej książki¹⁵. Na język polski została przetłumaczona jedna z najważniejszych książek Hare'a: *Myślenie moralne*. Wcześniej po polsku ukazały się cztery eseje: *Sąd moralny*, *Sprawiedliwość i równość*, *Gra w obiecywanie* i *Uniwersalny preskrytywizm*¹⁶ i jeden głos w dyskusji¹⁷. W niniejszej pracy będziemy się więc opierać prawie wyłącznie na tekstach anglojęzycznych. Większość z nich skupia się jednak na wczesnych książkach i artykułach Hare'a – często nie bierze się pod uwagę nowszych jego prac, które publikował aż do swojej śmierci (2002). Można spostrzec, iż wiele idei Hare'a

¹⁵ Zob. K Saja, *Internalizm motywacyjny R. M. Hare'a*, w: „Analiza i Egzystencja” (5) 2007, s. 179-202; tenże, *Typologia stanowisk metaetycznych a nonkognitywizm współczesny*, „Kwartalnik Filozoficzny” (1) 2008, s. 27-47; tenże, *Semantyka moralności a kryteria wyboru teorii metaetycznej*, „Etyka” (40) 2008.

¹⁶ R. M. Hare, *Sąd moralny*, „Etyka” (11) 1973; tenże, *Sprawiedliwość i równość*, „Etyka” (15) 1977; tenże, *Gra w obiecywanie*, „Etyka” (21) 1984, s. 151-164; tenże, *Uniwersalny preskrytywizm*, w: P. Singer (red.), *Przewodnik po etyce*, Książka i Wiedza, Warszawa 1998, s. 499-511.

¹⁷ Tenże, *Warunkowe i bezwarunkowe obowiązywanie norma moralnych*, „Etyka” (22) 1986.

oraz stawianych mu zarzutów pod wpływem tekstów wydanych pod koniec jego życia jawi się w innym, bardziej korzystnym dla niego świetle. Zawarta w tej pracy analiza oparta została zatem na wszystkich najważniejszych, z punktu widzenia naszego tematu, publikacjach. Z racji ograniczonej ilości literatury poświęconej dokonaniom Hare'a¹⁸ niniejsze studium w wielu fragmentach ma charakter autorski. Nie istnieje bowiem, jak się wydaje, żaden kanon interpretacyjny dotyczący jego dokonań filozoficznych. Analizując jego wypowiedzi, będziemy starać się je odczytywać jako spójną całość tworzącą swego rodzaju system. Proporcje między fragmentami, które jedynie referują i streszczają tok rozważań Hare'a, a tymi, które go analizują czy wskazują możliwe interpretacje, są inne w każdym z paragrafów.

Pomimo, iż Hare zręcznie odpowiadał na stawiane mu zarzuty i bronił swojej teorii do końca życia, z biegiem lat jego system uległ powolnym przeobrażeniom, do czego przyznawał się niechętnie. Na przestrzeni czasu nie zmienił on jednak swego stanowiska w podstawowych kwestiach charakteru języka i rozumowania moralnego. Ów rdzeń nazwać można racjonalnym nondeskrytywizmem. Hare miał mocne przekonanie, że pewne problemy rozwiązał definitywnie. Wyróżnić można jednak kilka elementów, które uległy zmianie. Z biegiem lat inaczej pojmował on prawdziwość sądów etycznych, złagodził swój internalizm oraz w odmienny sposób przedstawił owo „coś więcej”, co według nondeskrytywizmu oprócz warunków prawdziwości wyznacza znaczenie sądów etycznych¹⁹.

¹⁸ Autorowi niniejszej rozprawy udało się odnaleźć i uzyskać dostęp jedynie do jednej pozycji monograficznej poświęconej Richardowi Hare'owi: P. A. Adomankwah, *The Moral Philosophy of R. M. Hare. A Vindication of Utilitarianism?*, Peter Lang, New York 1997.

¹⁹ Zob. R. M. Hare, *Comments*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 201-204. Hare w tym miejscu wypisuje zmiany, które dokonały się w jego myśleniu na temat moralności.

oraz stawianych mu zarzutów pod wpływem tekstów wydanych pod koniec jego życia jawi się w innym, bardziej korzystnym dla niego świetle. Zawarta w tej pracy analiza oparta została zatem na wszystkich najważniejszych, z punktu widzenia naszego tematu, publikacjach. Z racji ograniczonej ilości literatury poświęconej dokonaniom Hare'a¹⁸ niniejsze studium w wielu fragmentach ma charakter autorski. Nie istnieje bowiem, jak się wydaje, żaden kanon interpretacyjny dotyczący jego dokonań filozoficznych. Analizując jego wypowiedzi, będziemy starać się je odczytywać jako spójną całość tworzącą swego rodzaju system. Proporcje między fragmentami, które jedynie referują i streszczają tok rozważań Hare'a, a tymi, które go analizują czy wskazują możliwe interpretacje, są inne w każdym z paragrafów.

Pomimo, iż Hare zręcznie odpowiadał na stawiane mu zarzuty i bronił swojej teorii do końca życia, z biegiem lat jego system uległ powolnym przeobrażeniom, do czego przyznawał się niechętnie. Na przestrzeni czasu nie zmienił on jednak swego stanowiska w podstawowych kwestiach charakteru języka i rozumowania moralnego. Ów rdzeń nazwać można racjonalnym nondeskrytywizmem. Hare miał mocne przekonanie, że pewne problemy rozwiązał definitywnie. Wyróżnić można jednak kilka elementów, które uległy zmianie. Z biegiem lat inaczej pojmował on prawdziwość sądów etycznych, złagodził swój internalizm oraz w odmienny sposób przedstawił owo „coś więcej”, co według nondeskrytywizmu oprócz warunków prawdziwości wyznacza znaczenie sądów etycznych¹⁹.

Część I Źródła etyki R. M. Hare'a

¹⁸ Autorowi niniejszej rozprawy udało się odnaleźć i uzyskać dostęp jedynie do jednej pozycji monograficznej poświęconej Richardowi Hare'owi: P. A. Adomankwah, *The Moral Philosophy of R. M. Hare. A Vindication of Utilitarianism?*, Peter Lang, New York 1997.

¹⁹ Zob. R. M. Hare, *Comments*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 201-204. Hare w tym miejscu wypisuje zmiany, które dokonały się w jego myśleniu na temat moralności.

Aby właściwie zinterpretować system etyczny Richarda Hare'a, należy umieścić jego teorię w kontekście historycznym. W części I przedstawimy pokrótce jego biografię, dorobek naukowy, piętno, jakie odcisnęli na nim poprzednicy oraz zdarzenia, które uformowały jego poglądy. Jego teoria nabrała kształtu pod wpływem: metaetycznych pytań postawionych przez G. E. Moore'a, emotywizmu, oksfordzkiej szkoły analizy języka potocznego oraz lektury I. Kanta i klasyków utilitaryzmu – J. Benthama, J. S. Milla i H. Sidgwicka. Ważnymi impulsami dla rozwoju zainteresowań Hare'a były również doświadczenia biedy końca lat dwudziestych XX w. oraz pobyt w obozie jeńnickim podczas II wojny światowej.

§1 Życie i dorobek R. M. Hare'a¹

Richard Mervyn Hare urodził się 21 marca 1919 roku w Backwell Down, dziesięć kilometrów na zachód od Bristolu. Był najmłodszym z pięciorga dzieci, synem Charlesa Francisa Aubone Hare'a

¹ Poniższy rys biograficzny został sporządzony na podstawie następujących materiałów: R. M. Hare, *A Philosophical Autobiography...*, dz. cyt.; J. Hare, *R. M. Hare: A Memorial Address*, „Utilitas” (3) 2002, s. 306-308; J. Griffin, *Richard Mervyn Hare*, „Journal of Applied Philosophy” (3) 2002; P. Singer, *R. M. Hare's Achievements...*, dz. cyt., s. 309-317; J. O'Grady, *Richard Hare*, „The Guardian”, 1 lutego 2002; C.C.W. Taylor, *Professor R. M. Hare*, „The Independent”, 6 luty 2002; Z. Szawarski, *Nekrologi: Richard Hare (1919 – 2002)*, w: *Polska Sieć Filozoficzna*. Pobrane z Internetu 15 października 2003 z <<http://www.psf.org.pl/publication.php?pid=327>>.

i Louise Kathleen Hare (Simonds). Rodzice R. M. Hare'a byli przedstawicielami wyższej klasy średniej, ojciec prowadził rodzinną firmę zajmującą się produkcją farb i powierzchni podłogowych. Rodzinny interes rozwijał się bardzo dobrze, aż do czasu wielkiej recesji końca trzeciej dekady ubiegłego wieku. W tym właśnie czasie, gdy Richard miał dziesięć lat, na zawał serca zmarł jego ojciec. Pogarszające się warunki materialne zmusiły rodzinę Hare'ów do obniżenia standardu życia. Matka, bez większych sukcesów próbująca utrzymać rodzinny interes, zmarła pięć lat później. Mając piętnaście lat, Hare pozostał bez rodziców. Następne lata swojego życia, aż do chwili poślubienia Catherine Verney, opisywał on jako „noc z przewagą sennych koszmarów”.

Od 1932 roku uczęszczał do szkoły w Rugby, gdzie zdobył przygotowanie z języków klasycznych. Po wielu latach wykorzystał je, pisząc monografię *Platon*². Dzięki uzyskanemu stypendium dalszą edukację kontynuował w Balliol College w Oksfordzie. Tam też po przerwie spowodowanej wojną w 1947 roku uzyskał tytuł magistra (MA). Następnie przejął obowiązki wykładowcy filozofii, które wykonywał do 1966 roku.

Po wybuchu kryzysu monachijskiego w 1938 roku Hare postanowił zdecydować czy zostać pacyfistą. Po napaści Niemiec na Polskę zgłosił się na ochotnika do Artylerii Królewskiej. Odrzucony początkowo z powodu słabego wzroku, dzięki swojej determinacji zdołał ostatecznie uzyskać zgodę na przyjęcie do armii. Jesienią 1940 roku został wysłany do Indii. W grudniu 1942, po upadku Singapuru, dostał się do niewoli japońskiej. W ciężkich warunkach spędził trzy i pół roku w więzieniu w Changi. Jak twierdził, spotykał tam sytuacje, w których wyniesione ze szkoły zasady moralne trudno było zastosować. Różnorodność kultur, którą dostrzegł, spowodowała, iż przestał wierzyć w uniwersalne i obiektywne standardy moralne poznawalne drogą intuicji. Narodziło się w nim głębokie przeświadczenie, artykułowane później we wszystkich jego książkach, że dobra teoria etyczna powinna przedstawić metodę podejmowania racjonalnych wyborów, które dotyczą nawet najtrudniejszych, niespotykanych dotychczas sytuacji.³ Młodzieńcze doświadczenie biedy, a następnie wojny i niewoli spowodowały, iż od

² R. M. Hare, *Platon*, Wydawnictwo Michał Urbański, Warszawa 1996.

³ Zob. tenże, *Why Do Applied Ethics*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 1-13.

etyki wymagał przede wszystkim praktycznego charakteru. Każdy system etyczny, który nie sprostałby różnorodności warunków, w jakich podejmować musimy decyzje, różnorodności podzielanych przez narody intuicji, należało zdaniem Hare'a odrzucić.

W 1966 roku przeniósł się z Balliol College do Corpus Christi College, gdzie uzyskał tytuł profesora filozofii moralnej. Jego poprzednikami byli J. L. Austin i W. C. Kneale. W Oksfordzie nauczał do 1983 roku. Następnie jako Graduate Research Professor of Philosophy wykładał na Uniwersytecie Florydzkim, dzieląc swój czas przez następne jedenaście lat pomiędzy mieszkaniem w Gainesville w Stanach Zjednoczonych a Ewelme w Oxfordshire. Przeszedł na emeryturę w 1994 roku. Wśród jego uczniów znalazły się takie postaci, jak: Brian McGuinness, John Lucas, Bernard Williams czy Peter Singer. Miał trzy córki – Bridget, Luise, Ellin – syna Johna i sześciornuków. Zmarł w Ewelme, Oxfordshire, 29 stycznia 2002 roku.

R. M. Hare rozpoczął swą karierę filozoficzną w latach czterdziestych, dokonując próby odrzucenia emotywizmu. Pierwszą książkę – *The Language of Morals* – w której zawarł idee rozwijane, objaśniane i bronił przez resztę swojego życia, opublikował w 1952 roku. Poświęcił ją głównie zagadnieniom metaetycznym. Tam też podał definicję znaczenia słów moralnych, na podstawie której próbował później skonstruować teorię moralnego rozumowania. Zainspirowany zarówno emotywistyczną rekonstrukcją języka moralnego, jak i swoimi wcześniejszymi studiami nad logiką imperatywów⁴, głosił, iż sądy moralne nie wyrażają stwierdzeń, lecz są zaleceniami (preskrypcjami), które podlegają prawom logiki. Tym jednak, co odróżnia je od imperatywów, jest własność uniwersalizowalności, dzięki której mogą być one przedmiotem racjonalnego namysłu. Idee te dały początek nowej teorii metaetycznej – uniwersalnemu preskrytywizmowi.

Jego druga książka – *Freedom and Reason*⁵ – poruszała zagadnienia etyki normatywnej, stanowiąc próbę pogodzenia emotywizmu i racjonalizmu, problemu ludzkiej wolności rozumianej jako możliwość realizowania własnych interesów z rozumem, który nakłada na nią ograniczenia. Hare nie zgadzał się bowiem z poglądem A. J. Ayera,

⁴ Zob. tenże, *Imperative Sentences*, w: tenże, *Practical Inferences*, The Macmillan Press, London 1971.

⁵ Tenże, *Freedom and Reason*, Oxford University Press, Oxford 1963.

według którego szukanie w filozofii rozstrzygnięć problemów moralnych należałoby uznać za pomyłkę⁶. Mocno wierzył, że filozofia może pomóc ludziom odpowiedzieć na pytanie „Co powinniśmy czynić?” i jako przykład rozważań z etyki praktycznej we *Freedom and Reason* zanalizował problem rasizmu. Po przedstawieniu w *The Language of Morals* formalnych cech języka moralnego, jak sam pisał, był „naprawdę zdziwiony, gdy opisany przez niego system [...] okazał się utylityzmem”⁷.

W 1981 roku Hare wydał trzecią książkę – *Myślenie moralne*⁸. Była ona próbą scalenia i uzupełnienia zagadnień metaetycznych i etycznych zawartych w jego poprzednich pracach. Zaprezentował w niej również nowe idee. Przede wszystkim rozróżnił dwie płaszczyzny rozumowania moralnego: intuicyjną i krytyczną, dzięki czemu starał się odeprzeć klasyczne zarzuty skierowane przeciwko utylityzmowi, który głosi, iż utylityzm stoi w sprzeczności z intuicjami moralnymi. Celem *Myślenia moralnego* było również naświetlenie najważniejszych elementów jego systemu ze szczególnym uwzględnieniem tych z nich, które budziły kontrowersje. Przedstawił tam analizę problemu fanatyzmu, amoralizmu i słabości woli. Można powiedzieć, iż wydanie tej pozycji zapoczątkowało dojrzały etap filozofii Hare'a.

Czwartą i ostatnią książkę – *Sorting Out Ethics*⁹ – Hare opublikował w 1997 roku. Ustosunkowawszy się w niej do istniejących w metaetyce podziałów, starał się obronić stanowisko, iż tylko teorie o cechach nondeskryptywistycznego racjonalizmu mają szansę na filozoficzne przetrwanie – żadne inne odmiany teorii etycznych nie spełniają ważnych dla tej dziedziny kryteriów.

Oprócz wymienionych pozycji Hare napisał również około stu siedemdziesięciu artykułów i recenzji, które w ostatnim czasie zostały zebrane w następujących wydaniach: *Practical Inferences; Essays on Philosophical Method; Applications of Moral Philosophy; Essays on the Moral Concepts; Essays in Ethical Theory; Essays on Political Morality; Essays on Religion and Education; Essays on Bio-*

⁶ Zob. A. J. Ayer, *Language, Truth and Logic*, Dover Publications, New York 1946, s. 112.

⁷ R. M. Hare, *Freedom...*, dz. cyt., s. 290.

⁸ Tenże, *Myślenie...*, dz. cyt.

⁹ Tenże, *Sorting Out Ethics*, Clarendon Press, Oxford 1997.

*ethics; Objective Prescriptions and Other Essays*¹⁰. Warto wspomnieć również o zbiorze artykułów poświęconych jego twórczości, *Hare and Critics. Essays on „Moral Thinking”*¹¹, w którym odpowiada on na zarzuty postawione mu przez najwybitniejszych przedstawicieli anglosaskiej filozofii moralnej, takich jak: W. D. Hudson, R. B. Brandt, W. K. Frankena, A. Gibbard, J. Griffin, J. C. Harsanyi, T. Nagel, D. Richards, T. M. Scanlon, P. Singer, J. O. Urmson, Z. Vendler i B. Williams.

¹⁰ Tenże, *Practical Inferences...*, dz. cyt.; tenże, *Essays on Philosophical Method*, The Macmillan Press, London 1971; tenże, *Essays on the Moral Concepts*, University of California Press, Berkeley and Los Angeles 1972; tenże, *Essays in Ethical Theory*, Clarendon Press, Oxford 1989; tenże, *Essays on Political Morality*, Oxford University Press, Oxford 1989; tenże, *Essays on Religion and Education*, Oxford University Press, Oxford 1992; tenże, *Essays on Bioethics*, Clarendon Press, Oxford 1993; tenże, *Objective Prescriptions and Other Essays*, Clarendon Press, Oxford 1999.

¹¹ D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt.

§ 2 Inspiracje filozoficzne

§ 2.1 G. E. Moore – błąd naturalistyczny

W 1903 roku w *Principia Ethica*¹ George Edward Moore wysunął brzemiennej w konsekwencje tezę: pojęcia moralne, takie jak np. „dobry”, nie mogą być definiowane za pomocą pojęć naturalnych². Zarzucił on niektórym teoriom etycznym, iż obarczone są błędem, nazwanym przez niego „naturalistycznym”³. Jako jeden z twórców filozofii analitycznej był przekonany, że problemy etyczne mogą być rozwiązane na drodze analizy języka. Jego zdaniem rozumieć znaczenie jakiegoś predykatu to tyle, co znać własność, do której się on odnosi. Aby zdefiniować dany termin, należy podać rozwinięte, syno-

¹ G. E. Moore, *Zasady etyki*, tłum. Cz. Znamierowski, De Agostini, Warszawa 2003.

² *Pojęcia naturalne* są pojęciami odnoszącymi się do własności naturalnych. „Przez naturę tedy rozumiem to, co jest przedmiotem badania nauk przyrodniczych i psychologii. Można powiedzieć, że natura obejmuje wszystko to, co istniało, co istnieje i co istnieć będzie w czasie. [...] Otóż postawiony przeze mnie problem istnienia w czasie stosuje się również i do cech. Czy możemy sobie wyobrazić, że cecha, którą oznaczamy terminem dobry, istnieje w czasie *samodzielnie*, nie zaś po prostu jako cecha przedmiotu naturalnego? Co do mnie, wyobrazić sobie tego nie mogę, gdy tymczasem w odniesieniu do bardzo znacznej ilości cech, tych właśnie, które nazywam cechami empirycznymi, da się według mnie pomyśleć, że istnieją one w czasie, niezależnie od istnienia innych przedmiotów.”; tamże, s. 63-64. Powyższe rozumienie własności naturalnych zostało skrytykowane. Zob. C. D. Broad, *Certain Features in Moore's Ethical Doctrines*, w: P. A. Schilpp (ed.), *The Philosophy of G. E. Moore*, Northwestern University, Evanston 1942, s. 41-68. Przez *predykat naturalny* rozumiemy będziemy takie słowo, którego prawdziwość przypisania go jakiemuś przedmiotowi może być rozstrzygnięta na drodze empirycznej. Spotkać możemy definicję mocniejszą: predykat naturalny jest takim predykatem, którego prawdziwość przypisania go danemu przedmiotowi może być zweryfikowana za pomocą metod nauk przyrodniczych.

³ Zob. G. E. Moore, *Zasady...*, dz. cyt., s. 16.

nimiczne wyrażenie, które można prawdziwie orzec o własności nazywanej przez ów termin. Okazuje się jednak, iż istnieją pojęcia, których zdefiniować nie można. Takim pojęciem jest „dobry”⁴. Moore jako dowód tej tezy zaprezentował tzw. **test otwartego pytania**. Można go zrekonstruować w następujący sposób⁵:

Załóżmy, że:

1) Predykat „dobry” jest analitycznym odpowiednikiem naturalnego predykatu „N”.

Zatem, na mocy definicji analityczności wiemy, że:

2) Znaczenie tezy „x jest N” jest identyczne ze znaczeniem tezy „x jest dobry”.

Gdyby tak jednak było, to:

3) Ktoś, kto poważnie zadaje pytanie „Czy x, który jest N, jest również dobry?”, popełniłby błąd pojęciowy.

Jednakże nie wydaje się, aby (3) było prawdą, gdyż,

4) Pytanie „Czy x, który jest N, jest również dobry?” jest *pytaniem otwartym* – znaczącym i niejawiającym się jako pomieszanie pojęć (np. można sensownie zapytać, „Czy czyn x, który jest *pożądanym przez większą część społeczeństwa*, jest również dobry?”).

Zatem:

5) Nie jest tak, że predykat „dobry” jest analitycznym odpowiednikiem predykatu „N”.

Zatem:

6) Na mocy konieczności pojęciowej własność *bycie dobrym* nie może być identyczna z własnością *bycie N*.

W konsekwencji Moore stwierdził, że własność dobra jest samostanna, prosta i niedefiniowalna, gdyż nie można zredukować jej do żadnego kompleksu własności złożonych. Jest ona również nienaturalna, gdyż istnieje poza czasem i przestrzenią⁶. Ponieważ dziedziną badania naukowego mogą być tylko cechy naturalne, naturalności⁷,

⁴ Zob. tamże, s. 10.

⁵ Por. A. Miller, *An Introduction...*, dz. cyt., s. 13-15.

⁶ Zob. G. E. Moore, *Zasady...*, dz. cyt., s. 64.

⁷ „Terminem więc «naturalizm» oznaczyłem pewną szczególną metodę stawiania zagadnień etycznych, metodę, która w istocie nie da się w ogóle pogodzić z możliwością istnienia etyki jako nauki. Metoda ta polega na zastąpieniu cechy «dobry» przez jakąś cechę empiryczną poszczególnej rzeczy natury lub zbioru rzeczy natury; w ten sposób zastępuje etykę przez jedną z nauk przyrodniczych”; tamże, s. 62.

zastępując przedmiot etyki przedmiotem jednej z nauk (najczęściej psychologii lub socjologii), postępują błędnie. Krytyka ta otworzyła nową kartę w dziejach metaetyki⁸.

Tezy Moore'a wywołały burzliwą reakcję i wzrost zainteresowania problematyką metaetyczną. Powstało kilka interpretacji *błędu naturalistycznego*.⁹ Według jednej z nich polega on na pomieszaniu lub utożsamianiu dwóch różnych własności o jakimkolwiek charakterze, a w konsekwencji definiowaniu jednej własności w terminach odnoszących się do drugiej. Według kolejnej, przez *błąd naturalistyczny* należy rozumieć jedynie próbę definiowania własności, które są z istoty swej niedefiniowalne (własności moralnych). W świetle następnej znanej interpretacji błąd ma miejsce gdy definiujemy predykaty moralne w terminologii pozaetycznej¹⁰.

Powyższy argument Moore'a stał się celem szerokiej krytyki. Przede wszystkim atakowano słabe uzasadnienie i kontrowersyjność założeń epistemologicznych i metodologicznych, które leżały u jego źródła¹¹. Współcześnie wielu filozofów twierdzi, iż Moore nie zdołał

wykazać błędności naturalizmu¹². Jednak pomimo tych zarzutów *test otwartego pytania* także dzisiaj w metaetyce ma duże znaczenie¹³. Dlaczego uważa się, iż *błąd naturalistyczny* niesie trafne intuicje, mimo że opiera się na wątpliwych przesłankach? Dla Hare'a najbardziej doniosłe wydaje się następujące wyjaśnienie¹⁴: pojęcia moralne, takie jak „dobro”, są powiązane z motywacją do działania. Pytania moralne dotyczą przede wszystkim tego, jak mamy postępować, zatem od odpowiedzi wymaga się, aby obejmowały motyw do działania. Pytanie Moore'a postawione w jego „teście” wydaje się otwarte, gdyż nasza wiedza dotycząca tego, co jest deskryptywnym, naturalistycznym synonimem pojęcia moralnego, nie pociąga za sobą motywu do działania. Dzieje się tak, ponieważ nie istnieje związek pomiędzy zdaniem opisowym a motywacją.

W ten sposób zinterpretowany przez Hare'a *test otwartego pytania* został przyjęty przez emotywistów¹⁵ i stał się jednym z głównych narzędzi krytyki deskryptywizmu. Wyjaśnienie powyższe pozwala również zrozumieć, dlaczego wywołał on tak znaczny odzew wśród metaetyków, nie wywierając żadnego wpływu na inne dziedziny filozofii, choć w wielu z nich redukcjonizm naturalistyczny jest stanowiskiem szeroko akceptowanym. Dzieje się tak, ponie-

⁸ Zob. S. Darwall, A. Gibbard, P. Railton, *Toward Fin de Siecle Ethics: Some Trends*, „The Philosophical Review” (1) 1992, s. 115. Zob. także A. Miller, *An Introduction...*, dz. cyt., s. 10-25; S. Jedynek, *Błąd naturalistyczny*, w: „Etyka” (2) 1967, s. 295. Obronę etyki Moore'a znaleźć można w: B. H. Regan, *How to Be a Moorean*, „Ethics” (3) 2003, s. 651-678.

⁹ Por. S. Jedynek, *Błąd...*, dz. cyt., s. 289-297, a także W. K. Frankena, *The Naturalistic Fallacy*, w: W. Sellars, J. Hospers (ed.), *Readings in Ethical Theory*, Appleton-Century-Crofts, New York 1952, s. 103-114.

¹⁰ Zdaniem W. K. Frankena „błąd naturalistyczny” nazwać należy „błędem definicyjnym”, polegającym na próbie sprowadzania pojęć moralnych do pojęć naturalnych; tamże. Jeśli jednak byłoby to prawdą, zarzut Moore'a dotyczyłby nie tyle wszystkich koncepcji naturalistycznych, ile jedynie jednej z jego odmian – **naturalizmu definicyjnego (analitycznego)**. Jak zobaczymy poniżej, taka interpretacja tezy Moore'a przyczyniła się do powstania naturalizmu syntetycznego; zob. § 3.3.2. Naturalizm definicyjny to pogląd, iż moralne własności są identyczne lub redukowalne do własności naturalnych na mocy definicji. Akceptując takie podejście, twierdzi się, że w definicji: *x jest dobre wtedy i tylko wtedy gdy x jest pożądaną przez większość część społeczeństwa pomiędzy definiensem a definiendum* zachodzi związek analityczny. Dzięki temu wiemy, iż dwie własności, które są odniesieniem definiowanych powyżej terminów, są z sobą tożsame. Por. A. Miller, *An Introduction...*, dz. cyt., s. 12.

¹¹ G. C. Field zwrócił uwagę na niewłaściwe, zbyt wąskie rozumienie przez Moore'a pojęcia definicji i przypisanie jej nieodpowiedniej roli w etyce. Według niego Moore niesłusznie utożsamia analityczność z tautologicznością: myli się,

gdy zakłada, że definicja nie może dostarczać nowych informacji. Przykłada on również zbyt wielką wagę do definicji proponowanych przez naturalistów, które stanowią jedynie przyczynki do dalszych badań, a nie skończone rozwiązania: dociekań etycznych nie rozpoczyna się bowiem od wprowadzania definicji, definicje co najwyżej mogą stanowić konkluzję takich dociekań; zob. G. C. Field, *The Place of Definition in Ethics*, w: W. Sellars, J. Hospers (ed.), *Readings...*, dz. cyt., s. 92-102. W. K. Frankena zauważył z kolei, że Moore w *teście otwartego pytania* utożsamia sens i referencję, które, jak wykazał Frege, należałoby traktować jako różne. Gdyby bowiem odróżnić owe pojęcia, test straciłby swą zasadność. Frankena stwierdził również, iż *błąd naturalistyczny* nie może stanowić instrumentu, za pomocą którego odrzucić można nierozwinięty jeszcze w pełni naturalizm. W przyszłości może zostać podana poprawna definicja naturalistyczna: test nie może tego wykluczyć. Opiera się on bowiem na współczesnych intuicjach językowych, które w przyszłości, właśnie za sprawą ewentualnego zwrotu naturalistycznego, mogą ulec zmianie. Moore wobec tego nieprawomocnie uogólnienia negatywne wyniki dotychczasowych rozwiązań naturalistów. Zob. W. K. Frankena, *The Naturalistic...*, dz. cyt., s. 103-114.

¹² Por. S. Darwall, A. Gibbard, P. Railton, *Toward...*, dz. cyt., s. 115.

¹³ Por. tamże, s. 117. Por. również A. Miller, *An Introduction...*, dz. cyt., s. 21-22.

¹⁴ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 84-85.

¹⁵ Zob. S. Darwall, A. Gibbard, P. Railton, *Toward...*, dz. cyt., s. 119-120.

waż definicje naturalistyczne, proponowane np. w epistemologii (definicja prawdy, wiedzy itd.) nie muszą zakładać związku pojęć z motywami do działania, nie muszą być traktowane jako normatywne.

§ 2.2 Emotywizm

Prawdziwym beneficjentem Moore'owskiego *testu otwartego pytania* oraz związanego z nim problemu przejścia od „jest” do „powinien”¹⁶ był zatem nonkognitywizm¹⁷. W połowie XX wieku stał się on najbardziej wpływowym rodzajem teorii metaetycznej. Pod koniec życia nawet Moore wahał się pomiędzy stanowiskiem Ch. L. Stevenson'a a kognitywizmem¹⁸.

Poglądy na temat moralności, które były podzielane przez emotywiistów, odnaleźć można już u Davida Hume'a.¹⁹ Argumentował on, iż moralność jest sprawą uczuć, a nie rozumu, zatem wypowiedane oceny są wyrazem przyjętych postaw lub odczuć. Oddzielił on rozum od woli, twierdząc, że jedynie uczucia mogą nas motywować: rozum nie może być nigdy motywem aktu woli ani przeciwstawić

¹⁶ Zob. D. Hume, *Traktat o naturze ludzkiej*, t. II, tłum. Cz. Znamierowski, Warszawa 1963, s. 259-260. Współcześnie nadaje się mu często nazwę „prawa Hume'a” bądź „gilotyny Hume'a”. Generalnie rzecz ujmując głosi ono, iż nie można poprawnie wywnioskować powinności z faktów, „powinien” z „jest”, tzn. nie można wywnioskować żadnej normatywnej konkluzji z przesłanek, które zawierają jedynie zdania opisowe. Problem ten został opisany bardziej szczegółowo w § 2.2.

¹⁷ Definicję nonkognitywizmu znaleźć można w § 3.2.

¹⁸ „Muszę raz jeszcze powiedzieć, że jestem skłonny myśleć, że «słuszny» we wszystkich etycznych użyciach, i oczywiście «niesłuszny», «powinien», «obowiązek» również nie są [...] w ogóle nazwami ocen, mają tylko znaczenie emotywne i nie posiadają w ogóle «znaczenia poznawczego»; i jeżeli jest to prawdą o tych słowach, to musi także być prawdą o «dobry», w tym sensie, który najczęściej rozważałem. Jestem skłonny myśleć, że jest tak, ale jestem także skłonny myśleć, że nie jest tak; i nie wiem, w którą stronę skłaniam się silniej.”; G. E. Moore, *A Reply to My Critics*, w: P. A. Schilpp (ed.), *The Philosophy of G. E. Moore*, Northwestern University Press, Evanston 1942, s. 554. Tłum. za B. Mejbaum, *Moore a Stevenson: spór o znaczenie terminów wartościujących*, „Etyka”, (10) 1972, s. 116. Analizę polemiki Moore'a ze Stevensonem znaleźć można tamże, s. 105-117.

¹⁹ Zob. Ch. L. Stevenson, *Ethics and Language*, Yale University Press, New Haven 1944, s. VII. Por. również M. Rutkowski, *Rola rozumu w decyzjach moralnych*. *Etyka Davida Hume'a*, Wyd. Uniwersytetu Wrocławskiego, Wrocław 201, s. 8, 149.

się uczuciom²⁰. „Rozum jest i winien być tylko niewolnikiem uczuć i nie może mieć nigdy roszczenia do innej funkcji niż do tego, żeby uczuciom służyć posłusznie”²¹. Z tego powodu, zdaniem Hume'a, geneza powinności tkwi w sferze emocji. Moralność wpływa bowiem na działania i uczucia, rozum zaś nie posiada takiej władzy.

Innym powodem powstania współczesnego emotywizmu były narodziny neopozytywizmu Koła Wiedeńskiego²². Ponieważ środowisko to rozpowszechniło bardzo restrykcyjne kryterium sensowności i znaczenia wyrażenia²³, co spowodowało wykluczenie etyki z grona „dobrych teorii naukowych”, należało zrozumieć na nowo, czym jest dyskurs etyczny oraz czym są jego pojęcia i sądy. Wyzwaniu temu próbowali sprostać tacy przedstawiciele emotywizmu jak: Axel Hägerström, Charles K. Ogden, Ivor A. Richards, Winston H. F. Barnes, Rudolf Carnap, Alfred J. Ayer i Charles L. Stevenson.

Emotywiizm podzielić można na skrajny i umiarkowany²⁴. Do miana pierwszego zaliczyć należy poglądy Rudolfa Carnapa i Alfreda Ayera. Zdaniem Carnapa, „nie jest w ogóle rzeczą możliwą formułowanie zdań, które wypowiedałyby sąd wartościujący”²⁵.

²⁰ Zob. D. Hume, *Traktat...*, dz. cyt., s. 185-188.

²¹ Tamże, s. 188.

²² Koło Wiedeńskie w 1924 założył M. Schlick (1889–1939). Jego śmierć, a także związane z II wojną światową rozproszenie intelektualistów austriackich, położyły kres działalności grupy. Członkami Koła byli G. Bergmann, R. Carnap, H. Feigl, O. Neurath i F. Waismann. Bliską więź z pracami Koła utrzymywał L. Wittgenstein; w latach 1927–1929 regularnie uczestniczył w zebraniach, a następnie pozostawał w kontakcie ze Schlickiem i Waismannem. Przedstawiciele Koła stworzyli tzw. pozytywizm logiczny (kierunek ten nazywano również empiryzmem logicznym i empiryzmem naukowym). W Anglii ruch ten zyskał popularność po ukazaniu się w 1936 r. dzieła Ayera *Language, Truth, and Logic*. Trwałym rezultatem oddziaływania tego nurtu jest szeroko rozpowszechniony brak zaufania do stwierdzeń, które nie mają żadnych kryteriów czy też warunków stwierdzalności: obiegowe hasło Wittgensteina, że znaczenie to sposób użycia, rozumie się często jako zlągodzony postulat respektowania zasady weryfikacji. Por. S. Blackburn, *Pozytywizm logiczny*, w: *Oksfordzki słownik filozoficzny*, J. Woleński (red.), Książka i Wiedza, Warszawa 1997, s. 305-306.

²³ Wyrażenie ma sens, jeśli istnieje kryterium jego weryfikacji (jeśli jest weryfikowalne). Znaczeniem wyrażenia jest metoda jego weryfikacji. Por. H. Buczyńska-Garewicz, *Koło Wiedeńskie*, Wyd. Comer, Toruń 1993, s. 48.

²⁴ Zob. M. Fritzhand, *Główne zagadnienia i kierunki metaetyki*, Książka i Wiedza, Warszawa 1970, s. 256-257.

²⁵ R. Carnap, *Oceny i normy to wypowiedzi pozbawione sensu*, w: I. Lazari-Pawłowska (red.), *Metaetyka*, PWN, Warszawa 1975, s. 82.

Wypowiedzi moralne to pseudozdania, które nie mają sensu. Nie mogą być więc ani prawdziwe, ani fałszywe. Pełnią one jednak swoją funkcję: „służą do wyrażania poczuc, poczuc związanych z życiem”²⁶. Linia demarkacyjna między tym, co jest, a co nie jest nauką, pokrywała się, według nich, z podziałem wypowiedzi na znaczące i te, które nie posiadają znaczenia. Najbardziej wyrazistą wersję takiej teorii przedstawił A. J. Ayer²⁷. Według niego zdanie „Kradzież pieniędzy jest zła” nie wyraża sądu, który mógłby być prawdziwy lub fałszywy, nie istnieją bowiem kryteria jego weryfikacji²⁸. Mówiąc, że działanie jest słuszne lub niesłuszne po prostu wyrażamy moralne uczucia²⁹. Zdania moralne nie są zatem twierdzeniami o naszych uczuciach, ale ich ekspresją, wykrzyknieniami w rodzaju: „ojej!”, „ach!”, „au!”. Przybierają one jedynie zwodniczą formę gramatyczną typową dla wypowiedzi oznajmujących.

Funkcja odnośnych słów etycznych jest czysto «emotywna». [...] Znaczenie różnych słów etycznych możemy definiować zarówno poprzez różne uczucia, które za ich pomocą wyrażamy, jak i poprzez powodowane przez nie reakcje, o których spodziewamy się, że wystąpią u odbiorców wypowiedzi³⁰.

Według Ayera, zdania etyczne nie mogą być ze sobą sprzeczne, relacja ta bowiem zachodzi jedynie między sądami logicznymi, które mogą być prawdziwe lub fałszywe³¹. Nie może zatem istnieć racjonalna polemika na temat wartości. Podczas dyskusji moralnej możemy mieć jedynie nadzieję, że „nasz przeciwnik zgodzi się z nami co do natury empirycznych faktów i przyjmie tę samą postawę wobec nich co my”³². Nie jesteśmy w stanie przytoczyć żadnych argumentów, które byłyby racją dla naszej postawy³³. Według emotywistów, dyskurs moralny jest bowiem wymianą perswazji a nie racji. Wypowiedzi moralne, kauzalnie zmieniając nasze psychiczne postawy, nie mogą być przesłankami ani wnioskiem żadnego argumentu.

Emotywizm został rozwinięty przez Charlesa L. Stevenson’a za sprawą wydanej w 1944 roku książki *Ethics and Language*³⁴. Na-

²⁶ Tamże, s. 83.

²⁷ A. J. Ayer, *Language...*, dz. cyt.

²⁸ Por. tamże, s. 107.

²⁹ Tamże.

³⁰ Tamże, s. 108.

³¹ Tamże, s. 109.

³² Tamże, s. 111.

³³ Tamże.

³⁴ Ch. L. Stevenson, *Ethics...*, dz. cyt.

stępnie Stevenson bronił go w wielu artykułach, które zostały zebrane w *Facts and Values* w 1963 roku³⁵. Choć akceptował większość założeń istotnych dla nonkognitywistów, teorii tej nadał bardziej umiarkowany charakter. Z poprzednikami łączyło go przekonanie o istnieniu zasadniczej różnicy pomiędzy funkcją języka nauki i etyki. Celem nauki jest poznanie, etyki zaś harmonizowanie ludzkich działań i postaw. W przeciwieństwie do Ayera Stevenson nie zaprzeczał jednak, iż wypowiedzi etyczne mogą być sensowne. Opierał się on bowiem na innej niż Ayer, tj. behawiorystycznej teorii znaczenia Charlesa W. Morrisa. Nazywana jest ona często również pragmatyczną lub psychologiczną teorią znaczenia³⁶. W podejściu tym „znaczenie» znaku musi być definiowane w terminach psychologicznych reakcji użytkowników tych znaków”³⁷ i jest jego własnością dyspozycyjną. Ponieważ istnieją różne rodzaje reakcji psychologicznych powodowanych przez dane wyrażenia, ich znaczenia można poklasyfikować. Dzięki temu Stevenson mógł wprowadzić fundamentalny dla Hare’a podział na znaczenie ematywne i deskryptywne; zauważył również, iż w jednej wypowiedzi mogą występować one wspólnie. „**Znaczenie ematywne** to takie znaczenie, w którym reakcja (z punktu widzenia odbiorcy) lub bodziec (z punktu widzenia nadawcy) ma charakter emocji”³⁸. Jest to taka dyspozycyjna własność zdania, która powoduje, iż nadaje się ono do wyrażania pewnych postaw osoby mówiącej (funkcja ekspresywna) lub do wywołania tych samych postaw u odbiorców (funkcja ewokatywna). Dzięki funkcji ewokatywnej wypowiedzi moralne to narzędzia, którymi możemy się posługiwać, aby zmieniać postawy innych. Za sprawą tej cechy teorię taką nazywa się często „instrumentalizmem”³⁹. Natomiast „**znaczenie deskryptywne** jest dyspozycją znaku do powodowania stanów poznawczych”⁴⁰. Jest to dyspozycyj-

³⁵ Tenże, *Facts and Values: Studies in Ethical Analysis*, Yale University Press, New Heaven 1963.

³⁶ Zob. Ch. L. Morris, *Foundations of the Theory of Signs*, w: *International Encyclopedia of Unified Science*, Vol. I, No. 2, University of Chicago Press, Chicago 1938.

³⁷ Ch. L. Stevenson, *Ethics...*, dz. cyt., s. 42.

³⁸ Tamże, s. 59.

³⁹ Zob. F. Ricken, *Etyka ogólna*, tłum. P. Domański, Wydawnictwo Antyk, Kęty 2001, s. 31.

⁴⁰ Ch. L. Stevenson, *Ethics...*, dz. cyt., s. 67.

na własność zdania, która odpowiada za wywoływanie postaw sądzieńowych (funkcja poznawcza) zarówno u odbiorcy jak i nadawcy zdania. Ponieważ znaczenie ematywne jest dyspozycyjną własnością zdania, nie każda wypowiedź o znaczeniu ematywnym musi być wyrazem emocji czy próbą zmienienia postaw innych. Stevenson uchronił się dzięki temu od subiektywizmu emotywnego skrajnego⁴¹. Pozostał jednak relatywistą kulturowym. Choć wypowiedzi moralne są znaczące, stwierdzenie, że jakiś znak coś oznacza, jest sensowne jedynie przy uwzględnieniu kontekstu jego użycia. W konsekwencji znaczenie jednobrzmiącej wypowiedzi może być różne w różnych grupach społecznych⁴². Aby adekwatnie oddać znaczenie wypowiedzi etycznej, należy wziąć pod uwagę dwa powyższe typy znaczeń, występują w niej one bowiem łącznie. Struktura głęboka wypowiedzi etycznych według Stevensona ma charakter zbliżony do poniższego schematu⁴³:

„To jest niesłuszne”	<i>Dezaprobuję to; czyń to również</i>
„On powinien to uczynić”	<i>Dezaprobuję poniechanie przezeń tego postępkę; czyń to również</i>
„To jest dobre”	<i>Aprobuję to; czyń to również</i>

Tabela 1. Ch. L. Stevensona schemat analizy

Pierwsza część zdania „Dezaprobuję to; czyń to również” jest wyrazem naszej postawy, a jej zgodność z prawdą może być sprawdzona empirycznie, druga natomiast nie posiada wartości logicznej, ponieważ jest imperatywem⁴⁴.

Zaprezentowane przez Stevensona dwie dystynkcje (imperatywne i propozycjonalne oraz ematywne i deskryptywne części znaczenia) stanowią klucz do zrozumienia metaetyki Hare’a. Richard Hare przejął bowiem niektóre idee emotywnego oraz wiążące się z nim zalety: emotywnizm w prosty sposób wyjaśnia występujące różnice

w ocenach tych samych czynów, odrębność języka normatywnego od opisowego, problemy kognitywistów z *testem otwartego pytania*, funkcję i genezę pytań etycznych, motywację do czynienia tego, co uważamy za słuszne, oraz globalną superweniencję predykatów moralnych na predykaty naturalnych, nie czyniąc przy tym żadnych metafizycznych założeń⁴⁵, które byłyby sprzeczne z naturalistycznym podejściem do świata⁴⁶. Emotywnizm posiada jednak również wady⁴⁷, które Hare starał się przezwyciężyć. Przede wszystkim, zaprzecza wielu fundamentalnym intuicjom lingwistycznym dotyczącym języka moralnego oraz prowadzi do relatywistycznych konsekwencji. W związku z ową sytuacją, zdaniem Hare’a, głównym celem metaetyki powinno być zasymilowanie pozytywnych aspektów emotywnizmu oraz uniknięcie jego kłopotliwych konsekwencji⁴⁸.

§ 2.3 Szkoła analizy języka potocznego

Jeśli metaetyczne problemy, z którymi borykał się Richard Hare, oraz cele, które sobie stawiał, wpływały ze splotu argumentów Moore’a i też emotywnizmu, to sposób ich podjęcia oraz język zaczerpnął on z oksfordzkiej szkoły filozofii analitycznej.

Przyjmuje się⁴⁹, iż filozofia analityczna zapoczątkowana została przez George’a Moore’a oraz Bertranda Russella w 1903 roku. W tym

⁴⁵ S. Blackburn, opierając się na prawie superweniencji, poddał krytyce realizm moralny. Według niego z założeń realistów oraz superweniencji wynika tajemnicza zależność: trwanie naturalnych własności logicznie gwarantuje ciągłość istnienia własności etycznych, choć własności naturalne, ponieważ realisci nie są redukcjonistami, logicznie nie gwarantują własności etycznych. Superweniencja dla realisty jest zatem mętnym, odizolowanym, logicznym prawem, dla którego nie może on przedstawić wyjaśnienia. Zob. S. Blackburn, *Supervenience Revisited*, w: G. Sayre-McCord (ed.), *Essays on Moral Realism*, Cornell University Press, Ithaca and London 1988, s. 64. Hare akceptował ten sposób krytyki realizmu: „Dr Blackburn, and I myself at one time, have thought that the supervenience of moral properties could be appealed to in order to refute such a view, and thus having knocked out both the naturalistic and the non-naturalistic varieties of realism, to establish anti-realism (or in my terms non-descriptivism)”; R. M. Hare, *Supervenience...*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 72.

⁴⁶ Naturalistyczne podejście do świata opisują reguły KH1–KH3, które znaleźć można w § 3.4.

⁴⁷ Zob. § 3.3.1.

⁴⁸ Zob. R. M. Hare, *Sorting...*, dz. cyt., s. 45-46, 126.

⁴⁹ Zob. M. Hempoliński, *Brytyjska filozofia analityczna*, Wiedza Powszechna, Warszawa 1974, s. 9.

⁴¹ Zob. tenże, *Ethics...*, dz. cyt., s. 54-56.

⁴² Zob. tamże, s. 56.

⁴³ Zob. tamże, s. 21.

⁴⁴ Tamże, s. 26.

właśnie czasie, choć data ta jest jedynie symboliczna⁵⁰, Moore na łamach czasopisma „Mind” opublikował esej *The Refutation of Idealism*⁵¹, Russell natomiast dzieło *The Principles of Mathematics*⁵². Choć prąd ten nie był homogeniczny, koncentrował się przede wszystkim na analizie pojęciowej, logicznej i lingwistycznej, przy czym różnice w sposobie analizy definiowały charakter każdej ze szkół. Dla filozofa tego nurtu jednym z ważniejszych narzędzi jest teoria znaczenia, która stanowi trzon działalności filozoficznej. Konkretne tezy często zależą właśnie od niej. Pierwszy ośrodek analizy powstał w Cambridge, gdzie dominowały dwie jej odmiany: analiza pojęciowa związana z osobą G. E. Moore’a oraz atomizm logiczny i lingwistyczny B. Russella i L. Wittgensteina. W latach dwudziestych w Wiedniu uformowała się kolejna grupa uczonych, skupionych wokół osoby Moritza Schlicka w Kole Wiedeńskim⁵³, która rozpoczęła następny etap jej rozwoju: logiczny pozytywizm. Po II Wojnie Światowej głównym ośrodkiem filozofii analitycznej stał się Uniwersytet Oksfordzki, gdzie pojawiła się czwarta odmiana analizy – filozofia języka potocznego, zwana również filozofią lingwistyczną. Jej twórcami byli: Gilbert Ryle, John L. Austin i Peter Strawson. W tym czasie również na Uniwersytecie w Cambridge dokonały się istotne zmiany wywołane nowym stanowiskiem L. Wittgensteina i wystąpieniem J. Wisdoma⁵⁴.

Richarda Hare’a zaliczyć należy do oksfordzkiej szkoły analizy języka potocznego. Jego filozoficzne przekonania najlepiej jest więc przedstawiać, odnosząc je m. in. do tez zaprezentowanych w *Dociekaniach filozoficznych* L. Wittgensteina, teorii aktów mowy J. Austina i performatywnej koncepcji prawdy P. Strawsona.

W latach trzydziestych Wittgenstein zaproponował nową metodę uprawiania filozofii: uważne badanie w języku naturalnym różnych sposobów użycia najbardziej interesujących filozofię wyrażen⁵⁵.

⁵⁰ Zob. T. Szubka, *Filozofia analityczna. Koncepcje, metody, ograniczenia*, książka nieopublikowana.

⁵¹ G. E. Moore, *The Refutation of Idealism*, „Mind” 12 (1903), s. 433-453.

⁵² B. Russell, *The Principles of Mathematics*, Cambridge University Press, Cambridge 1903.

⁵³ Zob. przypis 1 § 1.

⁵⁴ Por. M. Hempoliński, *Brytyjska...*, dz. cyt., s. 9-11.

⁵⁵ Zob. L. Wittgenstein, *Dociekania filozoficzne*, tłum. B. Wolniewicz, PWN, Warszawa 2000, s. 73-78.

Metoda ta polega na opisie sposobu użycia i funkcji słów lub powiązanych ze sobą całych ich zbiorów, nazywanych grami językowymi. „Gra językowa” (*language-game*) jest pojęciem należącym do późnej filozofii Ludwiga Wittgensteina. Gra językowa to wzorzec czynności i praktyk związany z jakąś szczególną rodziną wyrażen językowych. Termin ten wskazuje, iż posługiwanie się językiem powinniśmy traktować jako działalność podobną do gry rządzącej się własnymi regułami i w sobie zamkniętej⁵⁶. Gier takich może być nieskończenie wiele i są one od siebie niezależne. Opiera się na nich głęboki system gramatyczny języka. Termin „gra” ma według Wittgensteina podkreślać, że mówienie jest formą działalności, wyrazem pewnego sposobu życia. Wśród różnych form życia istnieją formy naukowe, religijne, estetyczne, etyczne, itd.⁵⁷ Według Wittgensteina w ramach kilku języków etnicznych, np. w angielskim i polskim, istnieje wiele „gier językowych”, dotyczących różnych dziedzin i poziomów życia. Każda z nich posiada inne reguły, które można opisać przez analogię do zasad rządzących grami uprawianymi w celach rozrywkowych, np. szachami czy piłką nożną. Aby poznać znaczenie wyrażenia, należy studiować sposób jego użycia w języku potocznym, sposób jego funkcjonowania w danej grze językowej. Powyższy zrąb teorii znaczenia wśród filozofów tego kręgu spowodował nastawienie anty-metafizyczne. Najmocniejszy temu wyraz dawał Gilbert Ryle⁵⁸, według którego głównym zadaniem analizy jest wykrywanie absurdalnych założeń filozoficznych i błędnych konstrukcji. Taki charakter ma większość rozwiązań filozoficznych, które uciekają się do wyjaśnień metafizycznych. Źródłem postulowania różnych bytów, takich jak substancjalnie odrębne umysły czy własności moralne, są błędy językowe.⁵⁹ Podobną postawę aprobował John Austin⁶⁰. Według niego

⁵⁶ Por. S. Blackburn, *Gra językowa*, w: tenże, *Oksfordzki...*, dz. cyt., s. 140.

⁵⁷ Por. A. Podsiad, *Gra językowa*, w: tenże, *Słownik terminów i pojęć filozoficznych*, Instytut Wydawniczy PAX, Warszawa 2000, s. 320.

⁵⁸ Zob. G. Ryle, *Czym jest umysł*, tłum. W. Marciszewski, PWN, Warszawa 1970.

⁵⁹ Zob. M. Hempoliński, *Brytyjska...*, dz. cyt., s. 93-94; G. Ryle, *Czym...*, dz. cyt., s. 37-40, 52-59.

⁶⁰ Zob. J. L. Austin, *Mówienie i poznawanie. Rozprawy i wykłady filozoficzne*, PWN, Warszawa 1993; J. Szymura, *Język, mowa i prawda w perspektywie fenomenologii lingwistycznej J. L. Austina*, Zakład Narodowy im. Ossolińskich – Wydawnictwo, Wrocław 1982.

zadanie filozofii polega na odkrywaniu rzeczywistej struktury i funkcji języka potocznego, w którym są formułowane najbardziej żywotne pytania filozofów. Wprowadzając nowy, sztuczny język, często uniemożliwiają oni udzielenie adekwatnych na nie odpowiedzi. Metodą rozwiązywania problemów filozoficznych jest więc analiza lingwistyczna. Z zaufaniem odnosi się ona do intuicji uczestników komunikacji, którzy od urodzenia mówią danym językiem⁶¹.

Richard Hare, jako przedstawiciel oksfordzkiej szkoły języka potocznego, akceptował zaprezentowany powyżej sposób filozofowania. Za L. Wittgensteinem traktował znaczenie jako użycie, interesując się przede wszystkim „grą językową moralności”. Zajmował się w szczególności użyciem takich potocznych terminów, jak: „powinien”, „musi”, „zły” i „dobry”. Opracował, krocząc podobną do Austina drogą, własny zarys teorii aktów mowy⁶², którym często posługiwał się w swoich tekstach, aby wyjaśnić funkcję wypowiedzi moralnych. Akceptował i powoływał się na performatywną teorię prawdy P. Strawsona⁶³. Sprzeciwiał się również rozważaniom metafizycznym, próbując, tak jak to postulował G. Ryle, problemy filozoficzne rozwiązywać na drodze analizy pojęciowej⁶⁴.

W związku z powyższą metodą filozofowania Bernard Williams postawił Hare’owi następujące pytanie: dlaczego studiowanie języka ma być bardziej wartościowe i uzasadnione niż studiowanie naszych intuicji moralnych?

Tym, którzy nie zgadzają się z Hare’em, że metoda lingwistyczna może stanowić fundament etyki, nie przedstawił on żadnych czysto metodologicznych racji, dlaczego nie mieliby oni zaczynać od «intuicji moralnych» – racji, które nie byłyby zależne od zajmowanych w sporze sprzecznych stanowisk⁶⁵.

Richard Hare wierzył jednak, że podstawowe pojęciowe prawa moralności znaleźć można w każdym języku używanym do dokonywania moralnych ocen⁶⁶. Co więcej, według niego rozpoczynanie dociekań etycznych od analizy znaczeń jest konieczne, ponieważ aby

⁶¹ Zob. J. Searle, *Czynności mowy. Rozważania z filozofii języka*, tłum. B. Chwedeńczuk, Instytut Wydawniczy PAX, Warszawa 1987, s. 28.

⁶² Zob. § 4.8.

⁶³ Zob. § 5.3.

⁶⁴ Zob. § 3.3.3.

⁶⁵ B. Williams, *The Structure of Hare’s Theory*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 195-196.

⁶⁶ Zob. R. M. Hare, *Sorting...*, dz. cyt., s. 129, 137, 141.

rozwiązać problemy moralne, trzeba zrozumieć pytania, które je wyrażają.

W filozofię moralności wkraczamy, zadając pewne pytania, one zaś zostają sformułowane w kategoriach pewnych pojęć. *Te właśnie* pojęcia dławią nas, gdy usiłujemy odpowiedzieć na owe pytania. [...] Dopóki nie poniechamy pytań, dopóty każda nowa forma słów, po które sięgamy, by wyrazić nasze pytania, musi zadowolić nas jako przeformułowanie *tych* pytań⁶⁷.

Jeśli celem dociekań etycznych są odpowiedzi na pytania dotyczące kierunku naszych działań, aby były one satysfakcjonujące, powinny być wyrażone w tym samym języku co pytania⁶⁸. Abyśmy mogli udzielić odpowiedzi, musimy więc znać znaczenia użytych pojęć oraz prawa logiczne dyskursu generującego problemy. Zatem analiza znaczeń języka moralnego stanowi konieczny warunek satysfakcjonującego rozwiązania nurtujących problemów etycznych⁶⁹. Według Hare’a, nie warto fundować teorii normatywnej na intuicjach niosących moralne treści, ponieważ ludziom łatwiej przychodzi celne rozstrzygnięcie o poprawnym czy niepoprawnym użyciu jakiegoś zwrotu niż o trafności danej intuicji moralnej⁷⁰. Jest empirycznym faktem, że intuicje lingwistyczne kompetentnych użytkowników danego języka są bardziej homogeniczne niż ich intuicje moralne. Przez **intuicje lingwistyczne** należy rozumieć intuicje dotyczące logiki danych gier językowych, których treścią są konieczne prawdy pojęciowe. Zdaniem Hare’a, są one czysto formalne i nie są, jak intuicje moralne, wypełnione materialnymi treściami⁷¹. Ludzie z reguły podobnie reagują na nielogiczne używanie języka, mają natomiast więcej wątpliwości, gdy rozważana jest słusność czy niesłusność jakiejś zasady etycznej⁷². Zaufanie do intuicji użytkowników, zdaniem Hare’a, powinno się ograniczać jedynie do dziedziny logiki czy lingwistyki⁷³. W szczególności nie ufał on ocenom wydanym przez intuicję, której przedmiotem były sądy ontologii, epi-

⁶⁷ Tenże, *Myślenie...*, dz. cyt., s. 31.

⁶⁸ Zob. tamże, s. 12-13.

⁶⁹ Zob. tamże, s. 197.

⁷⁰ Zob. tamże, s. 20, 23-24.

⁷¹ Zob. tamże, s. 20, 23. Jednak fakt, iż istnieje wiele poglądów dotyczących charakteru logiki wypowiedzi etycznych, czyni ten argument słabo uzasadnionym i nietrafnym.

⁷² Zob. tamże.

⁷³ Zob. tamże.

stemologii, psychologii czy etyki. Ich prawdziwość powinna zostać uzasadniona w niezależny od niej sposób.

§ 2.4 Kant i utylitaryzm

Zdaniem Hare'a, aby uniknąć irracjonalizmu i relatywizmu związanego z emotywizmem, przy konstruowaniu teorii etycznej należy wziąć pod uwagę cechy, które domenie moralności przypisał Immanuel Kant, i które prowadzą do specyficznej wersji utylitaryzmu. Etyka Hare'a jest bowiem połączeniem utylitaryzmu i kantyizmu, które scalił ze sobą, próbując uniknąć błędów swoich poprzedników i pozostawić to, co najbardziej wartościowe⁷⁴.

Można wyróżnić kilka elementów jego filozofii, które mają źródła u I. Kanta. Po pierwsze, obydwaj starali się podać metodę udzielania odpowiedzi na pytanie „Co powinniśmy zrobić?” bez zakładania jakiegokolwiek ujęcia dobra czy prawdziwości podzielanych przekonań moralnych. Wspólna była im próba ufundowania etyki normatywnej na własnościach formalnych. Zamiast rozpocząć refleksję od analizy danych empirycznych (zjawisk, naoczności) Kant, a za nim Hare, proponowali konstruowanie zasad moralnych na podstawie wymogów formalnych. W tym celu posługiwali się metodą uniwersalizacji⁷⁵, zakładającą, iż maksyma, która jest nieuogólnialna (nie można chcieć, aby obowiązywała powszechnie), musi zostać odrzucona. Zdaniem Kanta, właśnie taki charakter ma moralne prawo, a zasada, która się do niego odnosi, sformułowana jako żądanie, nazywa się imperatywem kategorycznym⁷⁶. Imperatyw Kanta: „Postępuj tylko według takiej maksymy, dzięki której możesz zarazem chcieć, żeby stała się powszechnym prawem”⁷⁷ jest podobny do tego, co stanowi treść myślenia moralnego według Hare'a, które w skrócie można sparafrazować jako imperatyw:

Jeśli chcesz postąpić moralnie, powinieneś postąpić zgodnie z taką zasadą wyrażającą zalecenie, o której zarazem chciałbyś z bezstronnej perspektywy, aby była uniwersalna.

⁷⁴ Zob. tenże, *Sorting...*, dz. cyt., s. 45-46, 126.

⁷⁵ Na temat uniwersalizacji Hare'a zob. § 6.

⁷⁶ „Kategorycznym byłby imperatyw, który przedstawiałby czyn jako sam w sobie, bez względu na inny cel, jako obiektywnie konieczny”; I. Kant, *Uzasadnienie metafizyki moralności*, przeł. M. Wartenberg, Antyk, Kęty 2002, s. 50.

⁷⁷ Tamże, s. 60.

Istotą kantyizmu Hare'a stanowi bowiem myśl, iż sądy moralne są podobne do imperatywów oraz mają charakter uniwersalny. Prowadzi to do pewnej wersji kantowskiego „państwa celów”, w myśl bowiem tej teorii moralność jest zbiorem praw ustanowionych przez ludzi, którzy są prawodawcami. Hare zgadzał się z Kantem, iż każda osoba jest prawodawcą, podlega prawu oraz jest autonomiczna – sama stanowi prawo pod warunkiem, że szanuje prawodawczość innych. Aby to uzasadnić, Kant posłużył się tradycyjną retoryką chrześcijańską i koncepcją umowy społecznej Rousseau⁷⁸, Hare natomiast analizą lingwistyczną oraz elementami utylitaryzmu.

Postulowana przez Hare'a uniwersalność i preskryptywność sądów moralnych powodują, że uniwersalny preskryptywizm uważać można za pewną odmianę kantyizmu. Potwierdza to również fakt, iż zarzuty stawiane Kantowi wykorzystywane były również przeciwko teorii Hare'a⁷⁹. Jednak Hare starał się uniknąć zastrzeżeń, które w stosunku do koncepcji jego poprzednika wydają się trafne. Dużo wysiłku włożył w ominięcie problemów związanych z deontologiczną naturą teorii Kanta. Wskazuje się bowiem, iż kantowskie maksymy mogą wchodzić ze sobą w konflikty, nie istnieje jednak ustalona procedura ich rozwiązywania⁸⁰. Hare starał się temu zaradzić, zwracając się w stronę konsekwencjalizmu i utylitaryzmu. Akceptował utylitaryzm, gdyż uważał, iż to skutki działania (przede wszystkim spełnienie bądź niespełnienie preferencji – interesów osób zależnych od danego działania) sprawiają, że czyn nazywamy słusznym bądź niesłusznym.⁸¹

Chcąc uprawomocnić próbę połączenia utylitaryzmu (konsekwencjalizmu) z kantyżmem, w ostatnim rozdziale *Sorting Out Ethics*⁸² starał się wykazać, iż można zasadnie postawić pytanie: „Czy Kant mógłby być utylitarystą?”. Choć był świadomy, że klasyczny utylitarysta J. Bentham i J. S. Mill jest sprzeczny z wieloma tezami Kan-

⁷⁸ Por. O. O'Neill, *Etyka Kantowska*, w: P. Singer (red.), *Przewodnik...*, s. 219.

⁷⁹ Wspólnie oskarża się ich np. o formalizm. Zarzut formalizmu głosi, iż imperatyw kategoryczny jest pusty i trywialny i nie określa żadnych zasad obowiązku. Zob. tamże, s. 221.

⁸⁰ Zob. tamże, s. 222.

⁸¹ R. M. Hare, *A Utilitarian Approach to Ethics*, w: tenże, *Objective Prescriptions...*, dz. cyt., s. 151.

⁸² Tenże, *Sorting...*, dz. cyt., s. 147-165.

ta, próbował dowieść, iż nie jest to cecha jego własnej odmiany utilitaryzmu. Chcąc „uwolnić Kanta od jego niektórych współczesnych uczniów”⁸³, tj. deontologów⁸⁴, tak jak J. S. Mill podał utilitarystyczną interpretację jego podstawowych tez. Formułę Celu Samego w Sobie⁸⁵ odczytał jako nakaz traktowania celów innych osób jako swoich własnych; Formułę Prawa Moralnego⁸⁶ zaś jako powinność chcenia tylko takich maksym, których stosowanie wpływa w najlepszy i bezstronny sposób na sytuacje osób zależnych od nas. Rezultatem jest utilitaryzm reguł, który nakazuje, aby każdy chciał tylko takich maksym, których konsekwencje oceniane z bezstronnego punktu widzenia najlepiej wpłyną na interesy wszystkich, których dana maksyma dotyczy⁸⁷. Jak potwierdza za J. S. Millem⁸⁸ O. O’Neill i czego bronił Hare, wbrew niektórym rozpowszechnionym interpretacjom etyki Kanta, jego zasada prawa powszechnego opiera się na przewidywanych skutkach działania zasad.

Gdy maksym nie da się uogólnić, to zazwyczaj dlatego, że zgoda na skutki ich powszechnego przyjęcia byłaby nie do pogodzenia ze zgodą na środki, jakich wymaga działanie na podstawie owych maksym⁸⁹.

Utilitarystyczne odczytanie etyki Kanta w wielu miejscach wydaje się jednak nadinterpretacją⁹⁰. Można bowiem zauważyć, iż

⁸³ Tamże, s. 149.

⁸⁴ Tamże, s. 152-154.

⁸⁵ „Postępuj tak, byś człowieczeństwa tak w twej osobie, jako też w osobie każdego innego używał zawsze zarazem jako celu, nigdy tylko jako środka”; I. Kant, *Uzasadnienie...*, dz. cyt., s. 73.

⁸⁶ Prawo moralne zdaniem Kanta wyraża imperatyw kategoryczny: „Postępuj tylko według takiej maksymy, dzięki której możesz zarazem chcieć, żeby stała się powszechnym prawem”; tamże, s. 60.

⁸⁷ R. M. Hare, *Sorting...*, dz. cyt., s. 153-154.

⁸⁸ „Gdy jednak zaczyna on [I. Kant – K. S.] wyprowadzać z tego przepisu [imperatywu kategorycznego – K. S.] któryś z konkretnych obowiązków moralnych, nie stara się – rzecz dziwna – nas przekonać, że przyjęcie przez wszystkie istoty rozumne jakichś najbardziej wyzywająco niemoralnych reguł postępowania zawierałoby jakąś sprzeczność, że byłoby logicznie (jeżeli nie fizycznie) niemożliwe; ogranicza się tylko do okazania, że gdyby tego rodzaju reguły zostały powszechnie przyjęte, pociągnęłyby to za sobą konsekwencje, na które nikt by się nie pisał”; J. S. Mill, *Utilitaryzm*, w: tenże, *Utilitaryzm. O wolności*, tłum. M. Ossowska, PWN, Warszawa 1959, s. 8.

⁸⁹ O. O’Neill, *Etyka...*, dz. cyt., s. 221.

⁹⁰ Na różnice między Kantem a Harem wskazywał H. M. Curtler w: tenże, *What Kant Might Say to Hare*, „Mind” (318) 1971, s. 295-297.

Kantowskie ujęcie uogólnialności różni się od pokrewnych (uniwersalny preskrytywizm, złota Reguła⁹¹) pod dwoma ważnymi względami. Po pierwsze, nie odwołuje się ono do tego, czego ktoś pragnie lub co preferuje, ani nawet do tego, że to, czego się pragnie lub co się preferuje, powinni robić wszyscy. Po drugie, procedura ta służy do wyróżniania tylko tych maksym, które należy odrzucić, jeżeli podstawowe zasady życia lub społeczeństwa mają być uogólnialne.⁹²

Wydaje się, iż przynajmniej w punkcie pierwszym opinia O. O’Neill trafnie wskazuje na różnice między Harem a Kantem⁹³. Kant oddzielił afekty, uczucia i związane z nimi pragnienia, w odniesieniu do których użył ogólnego terminu „skłonności”⁹⁴, od „wolnej woli”, rządzonej czystym rozumem. Jego zdaniem tylko zasady, które są wyrazem „woli”, mają charakter moralny, bo tylko wtedy będą one kategoryczne i nieprzygodne. Hare natomiast, choć dążył do zagwarantowania racjonalności dyskursu moralnego, nie odrzucił skłonności jako czynnika oceny. Jego zdaniem podstawą wydawania sądu moralnego jest każda preferencja, która może być uniwersalizowana, nawet ta o charakterze emocjonalnym. Z tego powodu zasady, które Hare nazywał moralnymi, Kant traktowałby jedynie jako reguły praktyczne⁹⁵ czy rady mądrości⁹⁶. Moralne zasady uniwersalnego preskrytywizmu mają bowiem formę uniwersalizowanego imperatywu hipotetycznego⁹⁷, a nie imperatywu kategorycznego.

⁹¹ Złotą regułę znaleźć można w pismach wielu religii. Chrześcijańska jej wersja brzmi następująco: „Wszystko co byście chcieli, żeby wam ludzie czynili i wy im czyńcie!”, Mt 7, 12.

⁹² O. O’Neill, *Etyka...*, dz. cyt., s. 222.

⁹³ Zob. § 6.2.

⁹⁴ „Zależność władzy pożądanego od uczuć (*Empfindung*) nazywa się skłonnością, która dlatego dowodzi zawsze jakiejś potrzeby.”; I. Kant, *Uzasadnienie...*, dz. cyt., s. 49.

⁹⁵ „Każdy musi przyznać, że prawo, jeżeli ma obowiązywać moralnie, tj. jako podstawa zobowiązania, musi się odznaczać absolutną koniecznością; [...] podstawy zobowiązania nie należy tutaj szukać w naturze człowieka ani w okolicznościach [zachodzących] w świecie, w jakim człowiek się znalazł, lecz *a priori* jedynie w pojęciach czystego rozumu, i że każdy inny przepis [...] jeżeli w najdrobniejszej części, być może nawet tylko co do pobudki, opiera się na empirycznych podstawach, może wprowadzić nazywać się praktyczną regułą, lecz nigdy prawem moralnym”; tamże, s. 12.

⁹⁶ Zob. tamże, s. 53.

⁹⁷ „Gdyby czyn był dobry tylko jako środek do *czegoś innego*, w takim razie imperatyw jest *hipotetyczny*; jeżeli przedstawimy sobie czyn jako dobry sam w sobie, a więc jako koniecznie [występujący] w woli zgodnej samej przez się z rozumem jako jej zasadą, to imperatyw jest *kategoryczny*”; tamże, s. 50.

nego w rozumieniu kantowskim. Co więcej, Hare, wartościując czynny ze względu na ich skutki, nie zwracał uwagi, w przeciwieństwie do Kanta, na motywy, które były ich przyczyną. Dla Kanta natomiast czyn jest moralnie dobry, jeśli został wykonany z obowiązku. „Albowiem jeżeli coś ma być moralnie dobre, to nie wystarcza, że *zgadza się* z prawem moralnym, lecz winno się także dokonać ze względu na nie”⁹⁸. Jedynym motywem, na którym można oprzeć konieczność zobowiązania, jest poszanowanie obowiązku. Nie akceptował on więc konsekwencjalizmu, którego Hare był zwolennikiem. Kant pisał o tym jednoznacznie:

Moralna wartość czynu nie leży w skutku, którego się po nim spodziewamy, a więc nie leży też w jakiegokolwiek zasadzie postępowania, która musi czerpać swą pobudkę z tego oczekiwanego skutku⁹⁹.

Według Kanta to, czy „możemy chcieć” obowiązywania danej maksymy, musimy stwierdzić, odwołując się do skutków powszechnego jej wprowadzenia, co nie znaczy, iż skutki czynu wyznaczają nam jego wartość. Hipotetyczne konsekwencje wprowadzenia maksymy mogą ją wykluczyć z grona reguł moralnych, jednak wartość czynu możemy ocenić, odwołując się jedynie do motywów.

Richard Hare podobieństw swojej teorii do etyki Kanta doszukiwał się również w tym, iż nie fundował on, tak jak i Kant, swojej teorii na przesłankach o treści materialnej. Według niego

Kant, jeśli dobrze go rozumiem, myślał że możemy ustanowić podwaliny czy fundament metafizyki moralności poprzez ujawnienie jej logicznej struktury. Ja także próbuję to uczynić¹⁰⁰.

Jednak oba projekty, wbrew pewnym podobieństwom, różnią się zasadniczo. Inaczej rozumie się w nich bowiem formalność. Dla Kanta wiedza formalna jest wiedzą dotyczącą form intelektu, bez uwzględnienia subiektywnych celów¹⁰¹, poznanie jest formalne, jeśli „zajmuje się tylko samą formą intelektu i rozumu oraz ogólnymi prawidłami myślenia w ogóle”¹⁰². Wiedza formalna dotyczy zatem

⁹⁸ Tamże, s. 13.

⁹⁹ Tamże, s. 30. Kanta krytykę utilitaryzmu znaleźć można również tamże, s. 15.

¹⁰⁰ R. M. Hare, *Fundamentalism and Coherentism*, w: tenże, *Objective Prescriptions...*, dz. cyt., s. 123.

¹⁰¹ Zdaniem Kanta: „Praktyczne zasady są *formalne*, jeżeli nie uwzględniają żadnych celów subiektywnych, są zaś *materialne*, jeżeli się na nich, a więc na pewnych pobudkach, opierają”; I. Kant, *Uzasadnienie...*, dz. cyt., s. 70.

¹⁰² Tamże, s. 9.

wszystkich podmiotów rozumnych. Hare natomiast przez formalność swoich tez pojmował własność bycia prawem pojęciowym obowiązującym w określonej grze językowej i nieprzyjmowanie żadnych materialnych założeń dotyczących np. natury ludzkiej, naszych potrzeb czy pragnień. W jego rozumieniu pojęcie formalności nie zakłada jednak traktowania pragnień jako nieistotnych przy ocenie moralnej. Powyższe rozumienie formalności stało się przedmiotem krytyki. Niektórzy uważali, że tak rozumiana formalność jest filozoficznie mało istotna, gdyż zrelatywizowana jest do języków etnicznych. Według A. MacIntyre’a

to, co Hare rozumie przez jedyny możliwy język moralności, wydaje mi się zaledwie jednym z kilku możliwych języków moralności. Dlatego też u podłoża wszelkich zastosowań języka moralności, które byłyby zgodne z Hare’a teorią znaczenia terminów moralnych, musi tkwić możliwość wyboru – zazwyczaj wyboru nieświadomego – umożliwiającego rozumienie moralności raczej w jakiś jeden określony sposób niż na wiele różnych, rywalizujących sposobów. Wybór ten zaś – ze względu na to, że poprzedza rozumowanie moralne, które zeń wynika – jest w tym sensie arbitralny¹⁰³.

Zarzut powyższy uderza jednak nie tylko w założenia analizy Hare’a, lecz w cały projekt szkoły analizy języka potocznego¹⁰⁴.

¹⁰³ A. MacIntyre, *Krótką historią etyki*, tłum. A. Chmielewski, PWN, Warszawa 2002, s. 19-20.

¹⁰⁴ Witold Marciszewski w wstępie do książki G. Ryle’a pisze: „Konfrontacja dwóch tylko języków, angielskiego i polskiego, dokonywana także przy okazji niniejszego przekładu, ukazuje na każdym kroku kruchość owej metody lingwistycznej w filozofii. Rzecz zaczyna się już przy próbie przekładu tytułowego terminu *mind*. Oczywiście, że przetłumaczymy to jako «umysł», bo co lepszego można zrobić? Ale przepadają w ten sposób ważne różnice znaczeniowe. Bo słowo *mind* jest jakby syntezą dwóch pojęć oddawanych po polsku słowami «dusza» oraz «psychika». Gdyby wartość książki Ryle’a miała polegać jedynie na tym, że dzięki analizie języka płoszy ona mity filozoficzne, to jej przekład na język polski byłby przedsięwzięciem mało celowym. Bo nikt po polsku nie popełni błędu pomieszania kategorii, mówiąc, że człowiek składa się z ciała i umysłu czy też z ciała i psychiki”; W. Marciszewski, *Wstęp. Gilbert Ryle i filozofia umysłu*, w: G. Ryle, *Czym...*, dz. cyt., s. 25-26.

Część II
Metaetyka

Dla prawdziwego filozofa chyba najbardziej ekscytującą rzeczą, lub może jedyną ekscytującą rzeczą, jest uzyskanie jasności co do pewnych ważnych pytań, które wcześniej były niejasne. Filozofią moralną zacząłem się ekscytować naprawdę wówczas, gdy jasno zacząłem postrzegać znaczenia słów moralnych takich jak „powinno się”. Fakt ten jest dla mnie poruszający częściowo ze względu na samo jego odkrycie, lecz także dlatego, iż jeśli tylko udałoby się wyjaśnić innym, co oznaczają te słowa i jaką ten fakt ma wagę dla naszych rozważań, mogłoby zostać rozwiązane wiele problemów na świecie¹.

§ 3 Nondeskryptywizm R. M. Hare'a

Uniwersalny preskryptywizm powszechnie zalicza się do nonkognitywizmu. Istnieją jednak mocne racje ku temu, aby porzucić tę klasyfikację. Sam Hare nazywał siebie nondeskryptywistą. Co więcej, uważał on, iż nondeskryptywizm jest najlepszą teorią metaetyczną². W niniejszym rozdziale przedstawimy Hare'a uzasadnienie powyższej oceny: zaprezentujemy akceptowaną przez niego koncepcję znaczenia i funkcji języka moralnego; alternatywne klasyfikacje teorii metaetycznych; jego stosunek do nonkognitywizmu, ekspresywizmu, emotywizmu, deskryptywizmu, naturalizmu, realizmu i intuicjonizmu; jego główne argumenty przeciwko konkurencyjnym stanowiskom metaetycznym oraz niektóre aspekty dyskusji wokół nondeskryptywizmu. Na koniec opiszemy zbiór kryteriów, którymi posługiwał się Hare, krytykując alternatywne koncepcje metaetyczne, oraz doda-

¹ R. M. Hare, *Philosophy and Conflict*, w: tenże, *Objective Prescriptions...*, dz. cyt., s. 56.

² Zob. tenże, *Sorting...*, s. 118, 136-137.

my do nich warunki, których wagę uznawali jego adwersarze. Utworzony w ten sposób zbiór kryteriów w kolejnych rozdziałach posłuży nam do oceny konkretnych tez uniwersalnego preskryptywizmu.

§ 3.1 Geneza i funkcja języka moralnego

John Austin zakładał, że wszystkie słowa posiadają rację dostateczną istnienia w języku: każde słowo ma w systemie, jaki stanowi język, określone miejsce i każde miejsce zajęte jest dokładnie przez jedno słowo³. Wydaje się, iż Hare podzielał tę opinię⁴. Według niego, nie tylko słowa mają swoje miejsce, także każda gra językowa ma swoją charakterystyczną funkcję. Żeby zrozumieć, co to jest młotek, nie wystarczy opisać jego struktury lub materiału, ale należy dowiedzieć się, do czego służy. Co więcej, to, czy dany młotek jest dobrym młotkiem, zależy od tego, czy dobrze spełnia on swoją funkcję. Zdaniem Hare'a podobna zależność zachodzi w dziedzinie etyki. Zrozumienie funkcji i celu moralnej gry językowej wydaje się być koniecznym warunkiem wypracowania dobrej teorii etycznej.

To, czemu służy język moralności, możemy zrozumieć, analizując wyrażone w nim pytania. Z badań nad nimi wynika, iż moralna gra językowa powstała, aby w ekonomiczny sposób przekazywać zalecenia, rady i nakazy⁵ oraz byśmy mogli postępować, nie wyrządzając sobie nawzajem krzywdy⁶. Język etyczny umożliwia dyskusję o naszych powinnościach i ograniczeniach naszego zachowania. Gdyby nie istniał, nie moglibyśmy ustalać reguł działania, które są odpowiedzialne za stabilną, przyczyniającą się do rozwoju ludzi koegzystencję⁷. Istnienie języka moralnego można wyjaśnić bez odwoływania się do racji moralnych, jego geneza ma bowiem swoje podłoże społeczne i wynika z potrzeby koordynacji ludzkich zachowań. Jest on swego rodzaju wydajnym narzędziem, które zostało stworzone, aby stanowić pomoc w rozwiązywaniu konfliktów pragnień i potrzeb⁸. Dlatego też w różnych społecznościach zaobserwować można istnie-

³ Zob. J. Szymura, *Język...*, dz. cyt., s. 137.

⁴ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 27.

⁵ Zob. tenże, *The Language...*, dz. cyt., s. 1; tenże, *Myślenie...*, dz. cyt., s. 94.

⁶ Tenże, *Objective Prescriptions*, w: tenże, *Objective Prescriptions...*, dz. cyt., s. 8.

⁷ Tamże, s. 9.

⁸ Zob. tenże, *Sorting...*, dz. cyt., s. 122-123.

nie odmiennych kryteriów oceny działań, rozwijały się one bowiem w odmiennych środowiskach (naturalnych, historycznych, kulturowych). Aby doprowadzić do lokalnej koordynacji zachowań, różne grupy ludzi wypracowały inne zasady moralne. Jednak *pierwotne znaczenie* sądów i podstawowych pojęć moralnych jest wszędzie takie samo⁹, gdyż wiąże się z uniwersalną funkcją, jaką one pełnią.

Hare nie zgodziłby się z postulatami przedstawianymi, np. przez Richarda Rorty'ego, iż jedynym rozwiązaniem mogącym rozwikłać problemy filozofii jest radykalna zmiana jej języka¹⁰. Według niego, tak jak zdaniem innych przedstawicieli oksfordzkiej szkoły analizy, język potoczny jest dobrym narzędziem i właściwie stosowany z powodzeniem spełnia swoją rolę. Jak pisał,

wszystkie moje poczynania w dziedzinie filozofii moralności [...] przekonały mnie, że pojęcia moralne, którymi powszechnie się posługujemy, dobrze nam służą, gdy już ustalimy ich uzus¹¹.

To brak ich zrozumienia powoduje widoczną we współczesnym świecie niemożność przezwyciężenia relatywizmu moralnego. Z tego względu fundamentalnym celem etyki jest objaśnianie znaczeń terminów i sądów jej języka¹². Musimy starać się odnaleźć logikę moralności, dzięki której na dalszym etapie moglibyśmy poznać reguły podejmowania moralnie słusznych wyborów. Aby napisać program komputerowy, który będzie zwyciężał w grze w szachy, konieczne jest poznanie reguł gry w szachy i reguł pisania programów. Analogicznie, aby zbudować dobrą teorię etyczną, należy poznać reguły języka moralnego, reguły racjonalności i reguły budowania dobrych teorii etycznych. Wiedza o nich pozwoli nam zrealizować pragnienie, by na terenie moralności nasze myślenie było bardziej racjonalne¹³.

§ 3.2 Klasyfikacja teorii etycznych

Klasyfikacja teorii filozoficznych, oprócz wartości porządkującej, posiada także inne ważne walory. Jej istotną cechą jest zwięzłe ku-

⁹ Zob. tamże, s. 129, 137, 141.

¹⁰ Zob. R. Rorty, *Nauka jako solidarność*, w: tenże, *Obiektywność, relatywizm i prawda. Pisma filozoficzne*, tom I, tłum. J. Margański, Fundacja Aletheia, Warszawa 1999, s. 56.

¹¹ R. M. Hare, *Myślenie...*, dz. cyt., s. 27.

¹² Zob. tenże, *The Language...*, dz. cyt., s. 1.

¹³ Zob. tenże, *Myślenie...*, dz. cyt., s. 11.

mulowanie informacji za pomocą pojęć. Informacje w ten sposób zmagazynowane ułatwiają filozofom poruszanie się w gąszczu teorii. Podobnie dzieje się w przypadku klasyfikacji stanowisk metaetycznych. Prowadzenie dyskusji na płaszczyźnie metaetycznej czy meta-metaetycznej często pozwala odnieść się do danego sposobu myślenia i pokazać jego słabości bez wchodzenia w rozważania przedmiotowe. Hare stosował tę metodę, krytykując nie konkretnych filozofów, ale np. naturalizm czy intuicjonizm. Z drugiej jednak strony, aby odrzucenie stanowiska metaetycznego pociągało za sobą zakwestionowanie sposobu myślenia pojedynczego filozofa, potrzebne jest właściwe zaklasyfikowanie jego poglądów. Często niestety zbyt pochopnie zaliczano uniwersalny preskrytywizm do takiej czy innej teorii metaetycznej. Przykładem tego jest wiele argumentów krytycznych rzekomo uderzających w preskrytywizm. Choć nierzadko dotyczą one emotywizmu czy nonkognitywizmu, nie czynią jednak szkody oryginalnej teorii Hare'a, gdyż wbrew panującym opiniom Hare nie był ani emotywistą¹⁴, ani nonkognitywistą. Teza ta będzie uzasadniana krok po kroku w poszczególnych rozdziałach niniejszej książki.

Standardowy sposób prezentacji teorii metaetycznych za pomocą dualizmu kognitywizm–nonkognitywizm, znaleźć można u Richarda Brandta. Teorię R. M. Hare'a zaliczył on do nonkognitywizmu¹⁵ (niekiedy zamiennie stosowane jest pojęcie akognitywizmu¹⁶). Tradycyjnie **nonkognitywizm** jest definiowany jako pogląd głoszący, iż wypowiedzi moralne nie są ani prawdziwe, ani fałszywe; ich wartość logiczna nie może być znana; nie odnoszą się one do faktów; nie wchodzą w skład ludzkiej wiedzy; nie są nawet przekonaniem, wyrażają bowiem inne stany umysłu, tj. emocje, preferencje lub nasze

¹⁴ Nieprawdziwą tezę, że Hare jest emotywistą, znaleźć można w pracach: A. Chmielewski, *Wprowadzenie. Filozofia moralności Alasdaira MacIntyre'a*, w: A. MacIntyre, *Dziedzictwo cnoty. Studium z teorii moralności*, tłum. A. Chmielewski, PWN, Warszawa 1996, s. XXII; tenże, *Niewspółmierność, nieprzekładalność, konflikt. Relatywizm we współczesnej filozofii analitycznej*, Wyd. Uniwersytetu Wrocławskiego, Wrocław 1997, s. 263; T. Ślipko, *Zarys etyki ogólnej*, Wydawnictwo WAM, Kraków 2002, s. 157; A. Podsiad, *Emotywizm*, w: tenże, *Słownik terminów i pojęć filozoficznych*, Pax, Warszawa 2000, s. 225.

¹⁵ R. B. Brandt, *Etyka...*, dz. cyt., s. 347-408.

¹⁶ Zob. P. Kawalec, *Akognitywizm*, w: Polskie Towarzystwo Tomasza z Akwinu, *Powszechna encyklopedia filozofii*, tom. I, Lublin 2000, s. 122-128.

zaangażowanie¹⁷. Do tak zdefiniowanego nonkognitywizmu zalicza się: emotywizm, preskrytywizm R. M. Hare'a, multifunkcjonalizm P. H. Nowell'a-Smitha¹⁸, projektywizm S. Blackburna¹⁹, norm-ekspresywizm A. Gibbarda²⁰ czy teorię kognitywistycznego ekspresywizmu M. Timmons²¹.

Współcześnie dualizm kognitywizm–nonkognitywizm ustępuje jednak miejsca nowej dychotomii, natomiast powyższe pojęcia uzyskały bardziej precyzyjne znaczenie. Aby móc porównać system Hare'a z koncepcjami najnowszymi, należałoby poczynić podobne przesunięcie pojęciowe. Nie tylko względy praktyczne uzasadniają zmianę terminologii. Również Hare przeciwstawiał się powyższemu sposobowi podziału teorii, nie bierze on bowiem pod uwagę złożoności uniwersalnego preskrytywizmu²². Choć nazywany nonkognitywistą, *explicitie* nie zgadzał się on z wieloma tezami wyznaczającymi znaczenie tego pojęcia. Twierdził np., iż

nondeskrytywista, taki jak ja, może zgodzić się, że sądy moralne mogą być prawdziwe lub fałszywe, że możemy wiedzieć o niektórych z nich, czy są prawdziwe, oraz uważać, że własności moralne istnieją.

Tradycyjna klasyfikacja przesądza natomiast z góry, iż żadna teoria należąca do nonkognitywizmu nie może spełnić wielu ważnych warunków, które wyznaczają zakres dobrych teorii etycznych: podatności na prawdę sądów etycznych, obiektywności wartości, istnienia racjonalnego dialogu etycznego, racjonalnej metody dochodzenia do wiedzy moralnej itd.²³ Pomimo że mianem nonkognitywizmu nazy-

¹⁷ Zob. Britannica Concise Encyclopedia, *Noncognitivism*. Encyclopaedia Britannica. Pobrane z Internetu dnia 4 wrzesień 2002 z <<http://concise.britannica.com/ebc/article?eu=398884>>; S. Blackburn, *Nonkognitywizm*, w: tenże, *Oksfordzki...*, dz. cyt., s. 261.

¹⁸ Według P. H. Nowella-Smitha słowa, które „filozofowie zwykli zwać «słowami wartościującymi», odgrywają wiele różnych ról. Używa się ich do wyrażania gustów i preferencji, do wyrażania decyzji i wyborów, do krytykowania, hierarchizowania, oceniania, doradzania, napominania, ostrzegania, perswadowania i wyperswadowania, chwaleń, ośmielania i onieśmiania, dekretowania i zwracania uwagi na reguły; bez wątpienia i do innych celów [...] byłoby absurdem zakładać, że [...] to samo wyrażenie może być używane tylko do jednej czynności”. Zob. P. H. Nowell-Smith, *Ethics*, Penguin Books, London 1965, s. 98.

¹⁹ Zob. przypis 5 we Wstępie.

²⁰ Zob. przypis 6 we Wstępie.

²¹ Zob. M. Timmons, *Morality Without Foundations...*, dz. cyt.

²² Zob. R. M. Hare, *Sorting...*, dz. cyt., s. 56.

²³ Por. tenże, *Ontology in Ethics*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 96.

wa się szeroki wachlarz systemów etycznych, to jednak wiele z nich nie jest zgodnych z większością cech przypisywanych nonkognitywizmowi (wszystkie warunki spełnia jedynie emotywizm).

Pojęcia nonkognitywizmu i kognitywizmu w standardowym ujęciu są nieadekwatne, ponieważ ich znaczenie wyznacza koniunkcja tez pochodzących z trzech niezależnych płaszczyzn: semantyki, epistemologii i metafizyki. W definicji nonkognitywizmu tezą semantyczną jest twierdzenie, że sądy etyczne nie wyrażają stwierdzeń, ale postawy od nich różne (np. emocje, zalecenia); epistemologiczną, że nie mogą być one prawdziwe lub fałszywe oraz że nie istnieje wiedza etyczna; metafizyczną zaś, że nie istnieją moralne fakty i własności. Niezależność powyższych trzech płaszczyzn względem siebie polega na tym, iż przyjęcie określonego stanowiska na gruncie jednej z nich (np. antyrealizmu) nie prowadzi do konieczności akceptacji konkretnego stanowiska przyjętego w ramach innej (np. ekspresywizmu). Z tego powodu zakres podpadających pod pojęcia kognitywizmu i nonkognitywizmu teorii jest bardzo wąski. Powiązanie ze sobą powyższych dziedzin jest spowodowane leżącą u źródeł standardowej klasyfikacji denotacyjno-konotacyjną teorią znaczenia oraz przekonaniem, które konstytuują wskazany przez J. L. Austina tzw. **błąd deskrytywizmu**²⁴. Zgodnie z tym sposobem myślenia wiedza dotyczyć może jedynie faktów, jest wyrażona poprzez sądy oznajmujące i składa się wyłącznie z przekonań. Tylko twierdzenia mogą posiadać wartości logiczne, prawda bowiem jest definiowana jako korespondencja sądu z faktami. W konsekwencji uważa się, że np. zwolennicy podejścia antyrealistycznego, którego Hare jest przedstawicielem, nie mogą posługiwać się kategorią prawdy. Stawiając taką tezę, nie bierze się jednak pod uwagę, iż we współczesnej epistemologii istnieją koncepcje, na gruncie których łączy się np. antyrealizm z podatnością na prawdę²⁵. Coraz większą rolę odgrywają również teorie etyczne, będące złożeniem niespotykanych wcześniej układów tez semantycznych, epistemologicznych i metafizycznych.²⁶

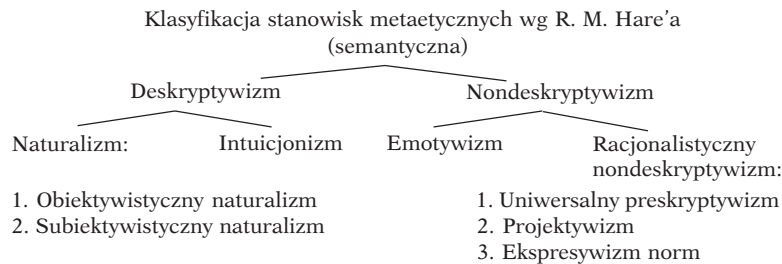
²⁴ J. L. Austin, *Mówienie...*, dz. cyt., s. 174, 552. Zob. § 3.3.2.

²⁵ Zob. 5.3.

²⁶ Ciekawym rozwiązaniem jest teoria błędu J. Mackiego, który łączy deskrytywizm z antyrealizmem. Zob. J. Mackie, *Ethics, Inventing Right and Wrong*, Penguin Books, New York 1977. Kolejnym interesującym systemem jest projektywizm S. Blackburna, który godzi nondeskrytywizm z quasi-realizmem. Zob. S. Black-

Istnieją trzy możliwości rozwiązania tego problemu: albo pozostawimy terminy „nonkognitywizm” i „kognitywizm” w użyciu, zdając sobie sprawę z wąskiego zakresu ich zastosowania, który w praktyce dla pierwszego z nich pokrywa się z zakresem użycia terminu „emotywizm”; albo zreformujemy ich znaczenie; albo przestaniemy ich używać i zastąpimy nowymi pojęciami. Richard Hare przyjął ostatnie rozwiązanie. Rzadko używał on pojęcia nonkognitywizmu, zastępując je terminem nondeskrytywizm²⁷. W tym miejscu nie zrezygnujemy jednak z używania tych rozpowszechnionych nazw, ponieważ po wprowadzeniu koniecznych uściśleń posłużą nam one do opisu pewnych koncepcji z dziedziny epistemologii moralnej.

Standardową dychotomię kognitywizm–nonkognitywizm najlepiej jest rozbić na trzy osobne podziały. Przebiegałyby one wzdłuż trzech płaszczyzn (semantycznej, epistemologicznej i metafizycznej), na których toczą się najważniejsze metaetyczne spory. W pracy *Sorting Out Ethics* Hare zaproponował podział teorii metaetycznych oparty na płaszczyźnie semantycznej. Traktował go jako aprioryczny, wyczerpujący oraz wolny od błędów modelu standardowego. Pytania, względem których został on ustalony, dotyczą takich kwestii, jak znaczenie sądów moralnych, ich lokucja, funkcja oraz moc illokucyjna. Przedstawia się on następująco²⁸:



Rysunek 1. Klasyfikacja stanowisk metaetycznych wg R. M. Hare'a

burn, *Spreading...*, dz. cyt. Również niestandardowym układem, charakteryzuje się teoria M. Timmons'a, w której autor broni fuzji nondeskrytywizmu i kognitywizmu. Zob. M. Timmons, *Morality Without Foundations...*, dz. cyt.

²⁷ Zob. R. M. Hare, *Objective...*, dz. cyt., s. 4.

²⁸ Zob. tenże, *Sorting...*, dz. cyt., s. 42. Na płaszczyźnie semantycznej konku-

1) **Deskrytywizm** – znaczenie sądów moralnych jest całkowicie zdeterminowane przez ich składnię i warunki prawdziwości²⁹.

2) **Nondeskrytywizm** – znaczenie oprócz składni i warunków prawdziwości zdeterminowane jest także czymś więcej.

Do deskrytywizmu (1) zaliczają się:

a) **Naturalizm** obiektywistyczny oraz subiektywistyczny – warunkami prawdziwości sądów moralnych są własności naturalne (subiektywne lub obiektywne).

b) **Intuicjonizm** – warunkami prawdziwości sądów moralnych są własności nienaturalne, „czysto moralne”.

Z kolei zwolennikami nondeskrytywizmu (2) są przedstawiciele:

a) **Emotywizmu** – sądy moralne nie podlegają prawom logiki.

b) **Racjonalistycznego nondeskrytywizmu** – sądy moralne podlegają prawom logiki. Do tej grupy Hare zalicza przedstawicieli uniwersalnego preskrytywizmu (ową logikę stanowią reguły rządzące uniwersalnymi preskrypcjami). Dodać tu można również projektywizm S. Blackburna, ekspresywizm norm A. Gibbarda i teorię M. Timmons'a.

rencyjną taksonomię wobec zaproponowanej przez Hare'a, stanowi dualizm ekspresywizm – deskrytywizm. To, co Hare nazwał nondeskrytywizmem, przyjęło się dziś nazywać ekspresywizmem. **Ekspresywizm** jest poglądem, który głosi, iż sądy moralne w swojej istocie nie są sędami opisowymi, ale pełnią zgoła inną funkcję: wyrażają postawy aprobaty i dezaprobaty (A. J. Ayer) lub bardziej skomplikowane dyspozycje do posiadania takich postaw (np. uniwersalne preferencje). Zob. M. Smith, *The Moral Problem*, Blackwell, Oxford 1994, s. 16. „Terminem tym określa się teorie dotyczące natury dyskursu etycznego, które przeciwstawiają wypowiedzi etyczne zdaniom wyrażającym przekonania. Teorie takie wśród funkcji, jakie pełnią wypowiedzi etyczne, pierwszeństwo przyznają albo wyrażaniu postaw, emocji i uprzedzeń, albo wydawaniu poleceń, albo wywieraniu presji celem nakłonienia do działania.”; S. Blackburn, *Ekspresywizm*, w: J. Woleński (red.), *Oksfordzki...*, dz. cyt., s. 104. „Expressivism denies that when we assert values, we talk about our own states of mind, in actual or potential circumstances. It says that we voice our states of mind, but denies that we thereby describe them.”; tenże, *Ruling Passions. A Theory of Practical Reasoning*, Clarendon Press, Oxford University Press, Oxford 1998, s. 50.

²⁹ Zauważyć trzeba, iż pojęcie „warunków prawdziwości” może być rozumiane w różny sposób, przede wszystkim ze względu na to czy przekraczają one ludzkie świadectwo, czy też mają do nich poznawczy dostęp. Hare jako antyrealista interpretuje je na sposób drugi. Według niego warunki prawdziwości wyznaczanie są społecznie. Szerzej ta kwestia omówiona została w § 4.3. W teorii Hare'a znaczenie deskrytywne wyznaczone jest przez warunki prawdziwości, znaczenie ewaluatywne zaś jest „czymś więcej” – konkretną mocą illokucyjną wypowiedzi.

Z przyjęcia propozycji Hare'a wypływają następujące korzyści: dzięki temu, że powyższe definicje są semantyczne, nie przemycają ukrytych założeń ontologicznych, żadnej definicji prawdy czy wiedzy; użyta do powyższych rozróżnień warunkowoprzawdziwościowa teoria znaczenia³⁰ nie została opisana szczegółowo, zatem to, jak rozumieć owe warunki prawdziwości, może być rozstrzygnięte podczas odrębnej dyskusji epistemologicznej lub zaproponowane przez samych twórców systemów etycznych; emotywizm A. J. Ayera i Ch. Stevensona, preskrytywizm R. M. Hare'a, projektywizm S. Blackburna, ekspresywnizm norm A. Gibbarda czy nondeskryptywistyczny kognitywizm M. Timmonsa zostały zaklasyfikowane, zgodnie z powszechną opinią, do jednego rodzaju teorii, tj. nondeskryptywizmu.

Na płaszczyźnie epistemologicznej najważniejsze pytania, które wpływają na podziały metaetyczne, brzmią następująco: (1) Czy sądy moralne mogą być prawdziwe/fałszywe?, (2) Czy wchodzą one w zakres naszej wiedzy?, (3) W jaki sposób poznajemy i uzasadniamy wartość moralną czynu? Jeśli znaczenie terminów kognitywizm i nonkognitywizm zostanie określone wyłącznie przez pojęcia teorii poznania, zgodnie z etymologią tych słów, to ich stosowanie wydaje się zasadne. Z tego powodu zbiór pozytywnych odpowiedzi na pierwsze i drugie pytanie nazwijmy kognitywizmem, negatywnych zaś – nonkognitywizmem. **Nonkognitywizm** byłby więc teorią, w ramach której sądy moralne nie przybierają wartości logicznych i nie wchodzą w skład ludzkiej wiedzy³¹. **Kognitywizm** by temu przeczył. Co więcej, w zależności od udzielonej odpowiedzi na trzecie pytanie kognitywizm moglibyśmy podzielić na słaby i mocny (wyznacza ona również to, czy daną koncepcję nazwać np. empiryzmem, racjonalizmem, naturalizmem, czy intuicjonizmem). Przez **kognitywizm mocny**, za A. Millerem³², rozumieć będziemy teorię, która twierdzi, iż (a) sądy moralne mogą być prawdziwe/fałszywe i (b) dokonywanie poprawnego (*correct*) osądu moralnego polega na ustanowieniu kontaktu poznawczego z niezależnie ukonstytuowanym stanem rze-

³⁰ Choć Hare nie używał tego pojęcia, trzeba zaznaczyć, iż z kontekstu wynika, że deskrytywizmowi przypisywał justyfikacyjną teorię znaczenia („rozumienie znaczenia zdania sprowadza się do znajomości warunków, w jakich usprawiedliwione będzie jego stwierdzenie”); T. Szubka, *Antyrealizm semantyczny. Studium analityczne*, Redakcja Wydawnictwa KUL, Lublin 2001, s. 77.

³¹ Zob. R. M. Hare, *Objective...*, dz. cyt., s. 4.

³² Zob. A. Miller, *An Introduction...*, dz. cyt., s. 138.

czy. **Kognitywizm słaby**³³ byłby natomiast teorią, w ramach której (a) sądy moralne mogą być prawdziwe/fałszywe i (b) dokonywanie poprawnego osądu moralnego nie polega na ustanowieniu kontaktu poznawczego z niezależnie ukonstytuowanym stanem rzeczy. Konsekwencją kognitywizmu słabego jest więc teza, iż własności moralne są zależne od naszych opinii, co najczęściej przyczynia się do akceptacji antyrealizmu.

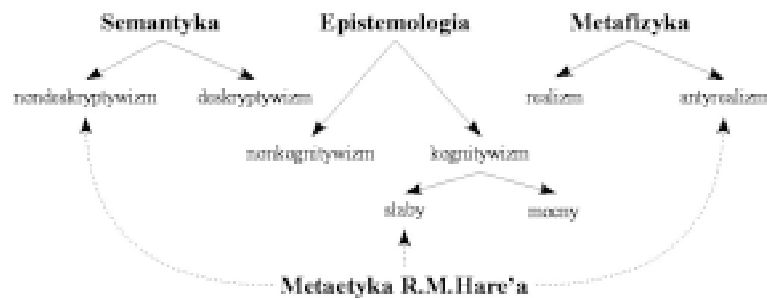
Trzecią płaszczyznę sporów stanowi metafizyka. Zakres zagadnień wyznaczony przez dociekania metafizyki moralnej konstytuują pytania dotyczące sposobu istnienia ocen i wartości, struktury i charakteru bytów. Pytania, które się stawia najczęściej, to: (1) Czy wartości/powinności/oceny/fakty moralne istnieją realnie? (realizm, antyrealizm); (2) Jeśli tak, czy istnieją w sposób charakterystyczny dla świata fizycznego/faktów empirycznych (fizykalizm, faktualizm, redukcjonizm), czy inaczej (antyredukcjonizm, platonizm)?; (3) Ile jest własności moralnych i czy można je do siebie zredukować (monizm, dualizm, pluralizm, redukcjonizm)? Teorię nazwiemy redukcyjną, jeśli głosi ona, że pojęcia moralne są analizowalne za pomocą terminów pozamoralnych (**redukcjonizm semantyczny**) lub twierdzi, iż własności moralne są identyczne z własnościami pozamoralnymi (**redukcjonizm metafizyczny**). **Antyredukcjonizm semantyczny** byłby stanowiskiem, według którego nie istnieje żadna interesująca analityczna relacja pomiędzy pojęciami moralnymi a pozamoralnymi – ich identyfikacja może mieć jedynie charakter syntetyczny. Zgodnie z tym poglądem zamiast o identyfikacji własności, powinno się raczej mówić o wielorakiej realizacji, ukonstytuowaniu czy superwencji własności moralnych nad własnościami naturalnymi. Redukcjonizm semantyczny w wersji naturalistycznej (pojęcia moralne redukują się do pojęć naturalnych) jest rozwijany między innymi przez takich filozofów, jak G. Harman, R. Brandt, D. Lewis, P. Railton, F. Jackson i P. Pettit. Nieredukcyjna odmiana naturalizmu jest akceptowana natomiast przez zwolenników naturalizmu syntetycznego (realistów z Cornell³⁴) oraz nonnaturalistów. **Nonna-**

³³ Tamże.

³⁴ Według przedstawicieli **realizmu z Cornell** semantyka podstawowych pojęć moralnych, takich jak „dobry”, jest podobna do semantyki pojęć rodzajów naturalnych takich jak „woda” czy „złoto”, która została przedstawiona przez H. Putnama i S. Kripkego. Tak jak zdaniem Putnama, terminy „woda” i „H₂O”

naturalizm jest stanowiskiem, które zakłada **antyredukcjonizm metafizyczny (realizm moralny)**. Według realistów własności moralne istnieją *sui generis* i nie są identyczne czy redukowalne do własności naturalnych, należą bowiem do „odrębnego porządku bytów”. Uważają oni, że moralne fakty i własności są niezależne od ludzkich poglądów. Głównymi przedstawicielami takiego stanowiska byli G. E. Moore, W. D. Ross, H. A. Prichard, a współcześnie J. McDowell i D. Wiggins³⁵. **Antyrealizm** jest zaprzeczeniem realizmu.

Klasyfikacja stanowisk metaetycznych, którą posłużymy się w tej rozprawie, przedstawia się więc następująco:



Rysunek 2. Wielopłaszczyznowa klasyfikacja stanowisk metaetycznych

Biorąc pod uwagę powyższy podział, teorię Hare'a należy zaliczyć do: racjonalistycznego nondeskrytywizmu, kognitywizmu sł-

nie są synonimami, pomimo że odnoszą się do jednej własności, zdaniem realistów z Cornell termin „dobry” nie jest synonimiczny z żadnym zbiorem predykatów naturalnych, choć odnosi się on do własności naturalnej. Identyczność własności nie musi bowiem pociągać za sobą synonimiczności (analityczności) pojęć. Definicje pojęć moralnych nie mają więc charakteru analitycznego, ale syntetyczny (stąd naturalizm syntetyczny) – znaczenie terminu „dobry” jest regulowane mocą konieczności metafizycznej, nie zaś pojęciowej. Najbardziej znanymi przedstawicielami nowej fali realizmu naturalistycznego są tacy filozofowie, jak David Brink, Richard Boyd, Nicholas Sturgeon, Geoffrey Sayre-McCord. Bardziej szczegółowo stanowisko to omówione zostało w: K. M. Saja, *Hare'a-Horgan-Timmons argument przeciw deskrytywizmowi*, w: „Diametros” (3) 2005.

³⁵ Por. A. Miller, *An Introduction...*, dz. cyt., s. 4-5.

bego (Hare uważał, że sądy moralne mogą być prawdziwe bądź fałszywe możliwa jest wiedza moralna, która jest zależna od naszych opinii³⁶), oraz antyrealizmu.

§ 3.3 Krytyka teorii konkurencyjnych

Richard Hare, starając się uzasadnić nadzieje pokładane w uniwersalnym preskrytywizmie, często czynił to negatywnie, dowodząc, iż konkurencyjne względem jego własnej teorii są uwikłane w wiele problemów, z którymi nie mogą sobie poradzić.

§ 3.3.1 Krytyka emotywizmu

Jak wspomnieliśmy powyżej, na preskrytywizm bardzo duży wpływ wywarł emotywizm. Hare włączył się w stworzony przez A. J. Ayera i Ch. L. Stevenson'a nurt krytyki deskrytywizmu oraz, za sprawą toczącej się dyskusji wokół ich teorii, uświadomił sobie popełnione przez nich błędy. Tworząc swoją teorię, starał się je ominąć. Z tego względu zrozumienie i ocena systemu Hare'a wymaga przynajmniej pobieżnego zapoznania się z historią krytyki emotywizmu³⁷.

Zasadniczym powodem, dla którego Hare nie akceptował emotywizmu, było przeciwstawianie się wypływającej ze stanowiska Ayera i Stevenson'a tezie, iż dyskurs moralny jest irracjonalny. Nie może bowiem uchodzić za racjonalne rozstrzygnięcie sporu, które nastąpiło wskutek oddziaływania Stevensonowskich „racji wspierających”³⁸. Racje wspierające (*supporting reasons*) nie zmierzają – jak racje logiczne – do przekonania o prawdzie lub fałszu danej oceny, lecz jedynie do wywołania jej aprobaty lub dezaprobaty. Nie są one racjami logicznymi, a tylko bodźcami: nie zachodzi między nimi żaden związek logiczny, a jedynie przyczynowo-skutkowy. Z tego względu sceptycyzm w kwestii racjonalności dyskursu etyki powo-

³⁶ Zob. 5.3.

³⁷ Argumenty krytyczne wobec emotywizmu znaleźć można w: R. M. Hare, *The Language...*, dz. cyt., s. 10-16; R. B. Brandt, *Etyka...*, dz. cyt., s. 385-396; P. Vardy, P. Grosch, *Etyka. Poglądy i problemy*, tłum. J. Łoziński, Wyd. Zysk i S-ka, Poznań 1995, s. 94; F. Ricken, *Etyka...*, dz. cyt., s. 40; A. MacIntyre, *Krótko...*, dz. cyt., s. 321-330; A. Miller, *An Introduction...*, dz. cyt., s. 38-43; M. Fritzhand, *Główne...*, dz. cyt., s. 263-301, 332-365.

³⁸ Zob. Ch. L. Stevenson, *Ethics...*, dz. cyt., s. 27-28.

duje konieczność zredukowania jej przedmiotu do psychologii. Co ważniejsze, postulowany irracjonalizm przeczy naszym fundamentalnym intuicjom dotyczącym sporów etycznych, dzięki którym mamy mocne przekonanie, że choć faktycznie dyskusje często przybierają formę perswazji i ekspresji emocji, to jednak w swym podstawowym wymiarze polegają na wymianie argumentów. Fakt, iż dyskusje etyczne często dzielimy na lepsze i gorsze z racji tego, czy prowadzone są oparte na perswazji czy raczej na argumentach o charakterze dowodów, świadczy, iż etyka racjonalna nie jest nam bynajmniej obca. Czy wewnętrzne roztrząsanie dylematów moralnych należałoby nazwać samoperswazją?

Nie może być również tak, że znaczenie sądów etycznych sprowadza się do wyrażania naszego nastawienia, często bowiem, gdy dokonujemy ocen etycznych, nie przyjmujemy żadnej postawy. Co więcej, zdarza się, iż nasze postawy są sprzeczne z sądami moralnymi, które akceptujemy. Można bez sprzeczności dać wyraz swemu upodobaniu, jednocześnie utrzymując, że jest to złe. Nieraz po dokonaniu refleksji moralnej dochodzimy do wniosku, że powinniśmy zaniechać tego, co najchętniej byśmy uczynili. Ktoś może bardzo lubić pornografię, może wykazywać sprzyjającą w stosunku do niej postawę, a zarazem uważać, że jest ona zła i powstrzymać się od jej oglądania. Ponieważ oceny moralne często przeciwstawiają się naszym uczuciom, praca moralistów, kaznodziejów i wychowawców polega na dostosowywaniu naszych postaw do sądów moralnych.

Jeśli natomiast nie każde uczucie przychylności jest postawą moralną, to jaki rodzaj uczuć wyrażają sądy moralne? Nie może być to zbiór charakterystyczny wyłącznie dla moralności, gdyż popełnialibyśmy wtedy błędne koło: sądy nazywalibyśmy moralnymi, jeśli wyrażałyby moralne uczucia, moralne uczucia zaś byłyby uczuciami, które są wyrażane przez moralne sądy³⁹. Z drugiej jednak strony, jeśli takie uczucia nie są specyficzne dla moralności, skąd będziemy wiedzieli, kiedy wyrażamy je w sensie moralnym? Czy istnieje więc intersubiektywnie dostępne kryterium rozstrzygnięcia tego, który sąd jest moralny?

Wypowiedzi etyczne nie mogą mieć wyłącznie znaczenia emotywnego, gdyż nie potrafilibyśmy odróżnić od siebie takich sądów, jak „Jan jest szlachetny” i „Jan jest dzielny”. To, że możemy je roz-

³⁹ Zob. A. MacIntyre, *Krótko...*, dz. cyt., s. 324.

różnić, spowodowane jest posiadaniem przez nie znaczenia deskryptywnego. Nawet jeśli weźmiemy pod uwagę umiarkowany emotytywizm Stevensona, w którym znaczenie sądów moralnych jest hybrydą dwóch elementów: znaczenia emotywnego i deskryptywnego⁴⁰, to również w stosunku do tej teorii wysunąć można zastrzeżenia. Nie jest przecież tak, że każda wypowiedź moralna jest połączeniem znaczenia emotywnego i deskryptywnego oraz że każda taka hybryda jest wypowiedzią moralną. „Gwałtu, rety, pali się!” jest niewątpliwie wypowiedzią hybrydową o znaczeniu zarówno deskryptywnym jak i emotywnym – nie jest to jednak wypowiedź etyczna. Sąd „Jan jest szlachetny” jest niewątpliwie wypowiedzią etyczną, nie musi jednak posiadać znaczenia emotywnego.

Emotytywizm jest nieadekwatny również dlatego, że będąc teorią subiektywistyczną, prowadzi do relatywizmu. W konsekwencji nie mielibyśmy możliwości obiektywnego i bezstronnego osądzania takich czynów jak ludobójstwo, ideologii jak nazizm czy ludzi jak Hitler. Nigdy nie moglibyśmy uznać za prawdziwe tego, że pewne czyny, ideologie czy ludzie są źli. To natomiast jest radykalnie sprzeczne zarówno z naszym doświadczeniem, jak i wymogami, które stawiamy wobec etyki. Nie spieramy się o prawdę czy fałsz wykrzykników, czynimy to jednak w odniesieniu do sądów etycznych. Sądy moralne, w przeciwieństwie do imperatywów czy ekspresji uczuć, posiadają bowiem wartości logiczne.

§ 3.3.2 Reductio ad absurdum deskrytywizmu⁴¹

Hare odrzucał nie tylko emotytywizm. Głównym zwalczanym przez niego stanowiskiem był deskrytywizm i naturalizm. W latach sześćdziesiątych XX wieku, pomimo krytyki naturalizmu dokonanej przez G. E. Moore'a, pojawiła się nowa fala deskryptywnego naturalizmu. Również na przełomie lat dziewięćdziesiątych XX wieku, za sprawą takich postaci, jak S. Blackburn czy A. Gibbard, odżyły dawne spory pomiędzy kognitywizmem a nonkognitywizmem. Obecnie wśród metaetyków jednym z intensywniej dyskutowanych argumentów przeciwko deskrytywizmowi jest tzw. argument *Moralnej Bliźnia-*

⁴⁰ Zob. Ch. L. Stevenson, *Ethics...*, dz. cyt., s. 71.

⁴¹ Rozdział ten w części pokrywa się z artykułem K. Saja, *Hare'a-Horgana-Timmons...*, dz. cyt., s. 56-74.

czej Ziemi Terrence'a Horgana i Marka Timmonsa. Zaprezentowany został on w serii artykułów opublikowanych na początku lat pięćdziesiątych⁴². Podobny argument przedstawił jednak dużo wcześniej Richard Hare'a, choć uczestnicy dyskusji, toczonej się wokół argumentu *Moralnej Bliźniaczej Ziemi* pomijają go milczeniem⁴³.

Po raz pierwszy ów zarzut Hare przedstawił na początku lat pięćdziesiątych w *The Language of Morals* pod postacią opowieści o *Misjonarzu i wyspie kanibali*⁴⁴, a następnie nadał mu nazwę *Reductio ad absurdum* deskrytywizmu w artykule wydanym w 1986 roku pod tym samym tytułem⁴⁵. Od tego czasu traktował go jako swoje podstawowe narzędzie krytyki deskrytywizmu⁴⁶. We wszystkich jego pracach, w których odwoływał się do tej argumentacji, stwierdzał, iż uderza ona w każdą odmianę deskrytywizmu.

Jego rekonstrukcję przedstawić można w następujący sposób⁴⁷:

(1) Jeśli znaczenie sądów, np. *Jedzenie ludzi jest złe* (p) i *Jedzenie ludzi nie jest złe* ($\sim p$), zależy tylko i wyłącznie od warunków prawdziwości, warunków stwierdzalności lub kryteriów poprawnego użycia słów (założmy deskrytywizm);

(2) i istnieją dwie odrębne społeczności, społeczność Misjonarzy (A) oraz Kanibali (B) gdzie w pierwszej sędzi się, że *Jedzenie ludzi jest złe* (p), w drugiej zaś, że *Jedzenie ludzi nie jest złe* ($\sim p$);

(3) oraz w społeczności Misjonarzy warunkami prawdziwości zda-

⁴² Zob. T. Horgan, M. Timmons, *New Wave Moral Realism Meets Moral Twin Earth*, „Journal of Philosophical Research” (16) 1991; tenże, *Troubles for new wave moral semantics: the 'open-question argument' revived*, „Philosophical Papers” (21) 1992; tenże, *Troubles on Moral Twin Earth: Moral Queerness Revived*, „Synthese” (92) 1992; tenże, *Copping Out On Moral Twin Earth*, „Synthese” (124) 2000; oraz M. Timmons, *Morality Without Foundations...*, dz. cyt.

⁴³ Szczegółowa analiza argumentu *Moralnej Bliźniaczej Ziemi* wraz z porównaniem go do omówionej poniżej krytyki deskrytywizmu dokonanej przez Hare'a została przeprowadzona w: K. Saja, *Hare'a-Horgana-Timmonsa...*, dz. cyt.

⁴⁴ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 148-149.

⁴⁵ Tenże, *A 'Reductio ad Absurdum' of Descriptivism*, w: tenże, *Essays in Ethical...*, dz. cyt.

⁴⁶ Argument ten znaleźć można również w wydanym po raz pierwszy w 1987r. artykule Hare'a *How to Decide Moral Question Rationally*. Tenże, *How to Decide Moral Question Rationally*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 102-103; oraz tenże, *Sorting...*, dz. cyt., s. 68-69.

⁴⁷ Rekonstrukcja opiera się przede wszystkim na: tenże, *The Language...*, dz. cyt., s. 148-149; tenże, *A 'Reductio...*, dz. cyt.; tenże, *Sorting...*, dz. cyt., s. 68-69; tenże, *Myslenie moralne*, s. 92.

nia *Jedzenie ludzi jest złe* (p) jest Q (np. fakt, iż jedzenie ludzi wymaga ich uśmiercania, co nie zgadza się z prawem naturalnym), a w społeczności B warunkami prawdziwości zdania *Jedzenie ludzi nie jest złe* ($\sim p$) jest W (np. fakt, że dusze osób zjedzonych osiągną wieczne szczęście w krainie śmierci);

(4) to powiedzieć trzeba, iż oba te sądy w różnych społecznościach mają różne znaczenia, które nie wchodzą ze sobą w relację sprzeczności. Społeczność A, stwierdzając: *Jedzenie ludzi jest złe*, mówi: *bowiem Zabijanie nie jest zgodne z prawem naturalnym* ($Q(a)$), a społeczność B, stwierdzając: *Jedzenie ludzi nie jest złe* ($\sim p$), mówi *Oso-by zjedzone zaznają wiecznego szczęścia* ($W(a)$). Ponieważ jest zarazem możliwe, że ($Q(a) \wedge W(a)$), możliwe jest również, że ($p \wedge \sim p$).

(5) Aby autentyczny spór moralny mógł zaistnieć, znaczenie terminów moralnych używanych przez członków różnych społeczności musi być takie samo.

(6) Zatem powiedzieć trzeba, iż na mocy przesłanek (1), (4) oraz (5) niemożliwy jest między nimi autentyczny dialog moralny; może zachodzić co najwyżej dyskusja językowa, spór o znaczenie⁴⁸.

(7) Jednak nasze intuicje lingwistyczne nie zgadzają się z powyższym wnioskiem (6):

(a) Nawet jeśli kryteria użycia danych ocen w różnych społecznościach są różne, nie znaczy to, iż nie może istnieć międzykulturowy, autentyczny konflikt przekonań moralnych. Może on istnieć i często ma miejsce. Co więcej, możliwe jest uczenie Kanibali etyki Misjonarzy i nie polegałoby to na uczeniu ich całkiem nowego języka. Kanibale rozumieliby, co mówiłby Misjonarz, nazywając coś dobrym, dziwiliby się jedynie, że nazywa on tak osoby, które są łagodne i nie kolekcjonują skalpów⁴⁹. „Prorocy i reformatorzy moralni mogą głosić kazania przeciwko nowym grzechom; mogą być uważani za szalonych, ale ich słowa będą zrozumiane”⁵⁰. Jednak w świetle deskrytywizmu spory moralne stają się czysto lingwistycznymi polemikami terminologicznymi. Ich treścią musi być znaczenie pojęć lub sądów,

⁴⁸ „Jeśli deskrytywizm byłby trafny, a słowo «zły» byłoby deskrytywne – czysto deskrytywne, nie mogłoby służyć do komunikacji”. Tenże, *A 'Reductio...*, dz. cyt., s. 117.

⁴⁹ Zob. tenże, *The Language...*, dz. cyt., s. 148.

⁵⁰ Tenże, *A 'Reductio...*, dz. cyt., s. 127.

a metodą osiągnięcia konkluzji przekonywanie swoich oponentów, że znaczenie ich jest takie, jak uważa jedna ze stron dyskusji.

(b) Gdyby znaczenie słów moralnych wyznaczone było jedynie przez standardy ich użycia, nawet jeśli znalazłbyśmy znaczenie wyrażenia „dobry kij do krykieta” czy „dobry chronometr”, nie potrafilibyśmy zastosować słowa „dobry” w odniesieniu do przedmiotów nieznanego nam wcześniej rodzaju. Możemy mówić jednak znacząco o „dobrej gaśnicy” czy „dobrym samochodzie”, nawet jeśli w stosunku do tych przedmiotów nie znamy kryteriów poprawnej aplikacji słowa „dobry”⁵¹.

(8) Co więcej, przyjęcie punktu (4) prowadzi deskrytywistów do bardzo niepożądanych konsekwencji. Analizując ewentualny spór wokół oceny walki na wojnie, Hare stwierdził w *Sorting Out Ethics*, że

walka może być zła w jednym, używanym przez jedną społeczność sensie słowa „zły” oraz nie być zła w drugim, używanym przez drugą społeczność sensie słowa „zły”. Jeśli ludzie w różnych społecznościach użyją swoich znaczeń słowa „zły” do oceny walki, ich oceny będą słuszne (*right*). [...] Naturalizm musi prowadzić do wniosku, iż zarazem jest słuszne w jednej kulturze nie potępiać walki oraz słuszne w drugiej kulturze ją potępiać. [...] Wobec tego zarówno ludzie, którzy nazywają akt walki złym, jak i ludzie, którzy nie nazywają go złym, mają równocześnie rację. *Mogliby* mieć równocześnie rację, jeśli słowo to znaczy różne rzeczy w dwóch kulturach. To natomiast jest pewnego rodzaju relatywizm⁵².

Wniosek taki jest *reductio ad absurdum* deskrytywizmu, gdyż cel, który chcą osiągnąć deskrytywiści (uzyskanie obiektywności wartości moralnych na racjonalnej i naukowej drodze), z konieczności musi być zaprzeczony⁵³.

(9) Najlepszą hipotezą wyjaśniającą powyższe problemy deskrytywizmu jest założenie, że sądy normatywne posiadają wspólny element preskryptywny, który determinuje ich znaczenie. Ujmowany jest on przez wszystkich użytkowników języka normatywnego nawet wówczas, gdy nie istnieje zgoda co do sądów opisowych dotyczących wartości czy racji uzasadniających stojących za danym

nakazem czy zakazem⁵⁴. *Reductio ad absurdum* jest więc argumentem na rzecz nondeskrytywizmu.

Podsumowując, powyższe argumenty Hare'a polegają na:

(I) Stwierdzeniu istnienia różnic w ocenach moralnych tych samych czynów w odmiennych kulturach. Hare zakłada (3) prawdziwość **deskrytywnego relatywizmu semantycznego**⁵⁵, który rozumiemy tutaj jako tezę empiryczną, iż w różnych społecznościach lub czasach obowiązują odmienne warunki prawdziwości, odmienne kryteria poprawnej aplikacji czy różne własności regulujące znaczenie terminów i sądów moralnych.

(II) Dowodzeniu, że na mocy deskrytywistycznej teorii znaczenia deskrytywizm z powyższego relatywizmu z konieczności prowadzi do (4) niewspółmierności schematów pojęciowych (**normatywnego relatywizmu semantycznego**⁵⁶).

(III) Założeniu prawdziwości intuicji lingwistycznej, która potwierdza autentyczność sporów moralnych w świecie realnym, pomimo występujących różnic w kryteriach poprawnej aplikacji pojęć moralnych.

(IV) Postawieniu wniosku, iż punkt (II), który jest konsekwencją deskrytywizmu, jest sprzeczny z naszymi intuicjami (III), nie może bowiem istnieć spór moralny, jeśli nie toczy się on przy użyciu pojęć o tych samych znaczeniach.

(V) Poparci wniosku stwierdzeniem, że w związku z (IV) deskrytywizm uniemożliwia zrealizowanie ambicji obiektywizacji ocen etycznych, co czyni go absurdalnym.

(VI) Wyprowadzeniu stąd ostatecznej tezy, że ponieważ należy przedkładać nasze intuicje lingwistyczne nad proponowane teorie znaczenia, jedynie nondeskrytywizm może poradzić sobie z owymi nieprzychylnymi dla deskrytywizmu konkluzjami. Tylko element

⁵⁴ Por. tamże, s. 119.

⁵⁵ *Deskrytywny relatywizm semantyczny* – teza empiryczna, która głosi, iż w różnych światach, społecznościach lub czasach ludzie mają odmienne przekonania na temat obowiązujących warunków prawdziwości, kryteriów poprawnej aplikacji czy własności regulujących znaczenie terminów i sądów. Zob. Ch. Swoyer, *Relativism*, w: E. N. Zalta (ed.), *The Stanford Encyclopedia...*, dz. cyt.

⁵⁶ *Normatywny relatywizm semantyczny* – znaczenie słów zdeterminowane jest przez społeczność językową, kulturę lub kontekst historyczny: dwie społeczności mogą posiadać wspólne słowo, choć będzie miało ono różne znaczenie. Zob. Ch. Swoyer, *Relativism...*, dz. cyt.

⁵¹ Por. tenże, *The Language...*, dz. cyt., s. 96, oraz tenże, *Myslenie...*, dz. cyt., s. 97.

⁵² Tamże, s. 69-70.

⁵³ Por. tenże, *A 'Reductio...*, dz. cyt., s. 113.

niedeskryptywny, który konstytuuje znaczenie, może być niezmienny i niezależny od akceptowanych warunków prawdziwości. Jedyne on umożliwia wyjaśnienie fenomenu sporu moralnego oraz daje nadzieję na obiektywizację ocen etycznych czy choćby osiągnięcie wzajemnego zrozumienia.

Nie wydaje się jednak, aby w starciu z powyższą krytyką deskryptywiści pozostali bezbronni. Istnieje kilka dróg wyjścia z impasu. Po pierwsze, jako naturaliści możemy nie zgodzić się na normatywny relatywizm semantyczny (niewspółmierność) (II), negując sposób opisu analizowanych sytuacji, a przede wszystkim ukryte założenie deskryptywnego relatywizmu semantycznego (I). Możemy twierdzić, iż przedstawiona argumentacja podlega błędnemu kołu, bowiem Hare interpretuje warunki prawdziwości w duchu antyrealizmu semantycznego⁵⁷: wykazana niewspółmierność języków społeczności A i B została wyprowadzona z ukrytego założenia, że znaczenia wyrażań to sposób ich użycia, a warunki prawdziwości są wyznaczane społecznie. Koncepcja znaczenia, którą przypisuje Hare deskryptywizmowi, nadaje znaczeniom charakter konwencjonalny. Nie musi być ona jednak podzielana przez realistów. Przez nich znaczenie jest bowiem często pojmowane jako odniesienie wyrażenia do obiektywnych stanów rzeczy, warunki prawdziwości jako przekraczające ludzkie świadectwo, terminy moralne zaś jako sztywne desygnatory. Wydaje się również, że nie ma podstaw, aby uważać – co w punkcie (8) czyni Hare – że przyjęcie perspektywy deskryptywizmu prowadzi od deskryptywnego relatywizmu semantycznego (3), poprzez pojęciową niewspółmierność sądów etycznych (4), do **normatywnego relatywizmu etycznego**⁵⁸ (8). Można bowiem twierdzić, że tylko jeden z powyższych sądów na temat walki na wojnie jest prawdziwy, tylko jeden z nich jest naprawdę słuszny – tylko jedna społeczność rzetelnie uchwyciła znaczenie. Druga popełniła błąd

⁵⁷ Antyrealizm semantyczny to stanowisko, które głosi, iż znaczenie zdań należy charakteryzować w kategoriach rozpoznawalnych warunków uprawomocnionej stwierdzalności; znaczeniem danego zdania są warunki pozwalające na jego uzasadnione przyjęcie. Por. P. Gutowski, T. Szubka (red.), *Filozofia brytyjska u schyłku XX wieku*, TN KUL, Lublin 1998, s. 60; T. Szubka, *Antyrealizm...*, dz. cyt., s. 62-91.

⁵⁸ *Normatywny relatywizm etyczny* – to co jest słuszne, sprawiedliwe czy cnotliwe jest względne wobec etycznych schematów pojęciowych, kontekstu kulturowego lub historycznego. Zob. Ch. Swoyer, *Relativism...*, dz. cyt.

poznawczy, gdyż zdaniu „Walka na wojnie jest zła” przypisała błędne warunki prawdziwości.

W konsekwencji jednak o ludziach, którzy nie uchwycili odpowiednich warunków prawdziwości lub nie odnoszą się do odpowiednich własności, a wydaje się, że takich osób jest wiele, powiedzieć należy, iż błędnie używają znaczenia sądów etycznych, posługują się nimi w błędnym znaczeniu. Rodzi się więc pytanie: czy sytuacja, która zakłada, iż jedna społeczność dobrze uchwyciła znaczenie, podczas gdy druga błędnie, nie prowadzi do konieczności stwierdzenia, że nie istnieje między nimi nawet spór językowy, a dyskusja jest nieporozumieniem? Akceptując powyższą obronę deskryptywizmu, znów stajemy przed problemem wyjaśnienia fenomenu sporu etycznego. Owszem, metaetycy czy naukowcy mogą dyskutować na temat tego, czy własność N odpowiada za znaczenie terminu „dobry”, czy nie, jest to jednak spór drugiego rzędu – spór dotyczący znaczeń, nie zaś normatywna dyskusja moralna, w której uczestniczymy codziennie.

Powyższa odpowiedź zakłada jednak poprawność i filozoficzną ważność intuicji lingwistycznych. Jako naturaliści możemy się na to nie zgodzić. Analizowane argumenty poddać można bowiem krytyce, uderzając w przesłankę (7)⁵⁹: intuicje semantyczne, które są jej źródłem, nie są konkluzywne. To, czy w przyszłości zostanie podana poprawna definicja naturalistyczna, nie może przecież w konieczny sposób zostać zanegowane przez nasze intuicje lingwistyczne lub test otwartego pytania. W przyszłości, za sprawą dokonanego zwrotu naturalistycznego, intuicje mogą ulec zmianie. Opieranie się na współczesnych intuicjach z góry zakłada, iż rewizjonizm naturalistyczny jest nietrafny. To, że dla Hare'a intuicje językowe są mocnym argumentem empirycznym, nie musi nas przekonywać. Kompetencja lingwistyczna rozpoznaje bowiem tylko aktualne normy semantyczne, te natomiast ulegają zmianom w miarę rozwoju nauk. Co więcej, relatywizm pojęciowy prowadzi do zanegowania dialogu, jeśli dwóch różnych języków nie można nawzajem przetłumaczyć. Naturalista może jednak uważać, że można odkryć język, który pozwoli nam przełożyć rzekomo niewspółmierne dyskursy obu społeczeństw. Dlaczego nie możemy wierzyć, że tak jak odkrycie pierwiastka złota rozwiązało problem znaczenia pojęcia „złoto”, tak

⁵⁹ Zob. A. Miller, *An Introduction...*, dz. cyt., s. 167-168.

samo znalezienie własności czy zbioru własności naturalnych rozwiąże problem znaczenia pojęć moralnych?

Wysunąć można jednak zastrzeżenie, że nawet jeśli udało się przekonać wszystkich do zmiany języka, ciągle pozostaje pytaniem otwartym, jak wyjaśnić fenomen sporu moralnego, który odbywa się tutaj i teraz, gdy nie znamy jeszcze owego uniwersalnego języka moralności. Konflikty musimy rozwiązywać już teraz, już teraz musimy prowadzić dialog moralny, deskrytywizm natomiast to nie-możliwia. Po drugie, nie należy być sceptykiem, jeśli nie ma się ku temu mocnych racji. Owszem, nasze intuicje językowe mogą być błędne, ale wszyscy, którzy tak uważają, muszą podać pewne wyjaśnienia⁶⁰: dlaczego są one tak mocne i szeroko rozpowszechnione, skoro są niepoprawne? Zatem choć intuicje mogą być błędne, nie mamy wystarczających powodów, aby nie dawać im wiary i nie traktować jako wyrazu kompetencji semantycznej, która odkrywa przed nami autentyczne własności języka.

Kolejnym możliwym sposobem odpowiedzi deskrytywisty na powyższe argumenty jest zanegowanie tezy o niemożliwości istnienia autentycznego sporu pomiędzy ludźmi, którzy używają słów o różnych znaczeniach (odrzuć przesłankę (5)). Czy koniecznie bowiem musimy twierdzić, że relatywizm pojęciowy prowadzi do zaprzeczenia istnienia dialogu? Istnieje metoda obrony możliwości dialogu, która, jak się wydaje, rozwiązuje powyższe problemy deskrytywizmu. Polega ona na rozróżnieniu pomiędzy odniesieniem nadawcy i odniesieniem semantycznym, między tym, do czego dany termin jest używany a tym, co sam znaczy. W ten sposób atakuje argument *Moralnej Bliźniaczej Ziemi* Heimir Geirsson⁶¹, krytyka ta ma zastosowanie jednak również do *reductio ad absurdum* deskrytywizmu Hare'a. Według Geirssona, aby zaistniała możliwość niezgodności dwóch wypowiedzi, nie jest konieczne, by terminy w nich zawarte miały wspólną ekstensję czy znaczenie. Wystarczy, żeby osoby uczestniczące w sporze odnosiły się za ich pomocą do tej samej rzeczy. Odniesienie nadawcy – odniesienie, na które intencjonalnie nakierowuje się

⁶⁰ Odpowiedź taką można znaleźć w: M. Timmons, *Morality Without Foundations...*, dz. cyt., s. 66.

⁶¹ Zob. H. Geirsson, *Moral Twin-Earth and Moral Semantic Realism*, „Erkenntnis” (62) 2005, s. 353-378.

nadawca danej wypowiedzi, które ma on na myśli, wypowiadając dane słowo – pozwala mu mówić o rzeczach, do których terminy same w sobie się nie odnoszą. Odniesienie nadawcy nie jest bowiem jednoznacznie zdeterminowane ekstensją czy znaczeniem samego wyrazu, ale również intencją, z jaką wyraz został wypowiedziany. Świadczy o tym np. metaforyczne lub ironiczne użycie języka. Co więcej, ludzie mogą się porozumiewać pomimo tego, że błędnie stosują słowa, których używają. Wyobraźmy sobie sytuację, w której dziecko, zawsze wskazując na tramwaj, używa określenia „lokomotywa”. Jeśli jego matka wie, że syn, wypowiadając słowo „lokomotywa”, konsekwentnie odnosi je do tramwaju, może prowadzić z nim sensowną konwersację na temat tramwajów za pomocą wyrazu „lokomotywa”.

Wydaje się, że błąd argumentu Hare'a tkwi w założeniu prawdziwości przesłanki (5). Polega on bowiem na wyprowadzaniu niekorzystnych wniosków pragmatycznych (negowaniu możliwości komunikacji) z konkretnej semantyki proponowanej przez deskrytywistów. Semantyka i pragmatyka nie muszą jednak być ze sobą powiązane w ten sposób, iż semantyczna niewspółmierność dwóch języków prowadzi do niemożliwości zaistnienia komunikacji.

Powiedzieć więc można, iż *reductio ad absurdum* deskrytywizmu, a w szczególności punkt (8), pozostawia nietkniętym deskrytywistyczne odmiany mocnego kognitywizmu, godząc jedynie w kognitywizm słaby. Co więcej, jeśli zgodzimy się z tezą, że niewspółmierność znaczeniowa nie musi powodować niemożności komunikacji, również przedstawiciele słabego kognitywizmu mogą wyjść obronną ręką.

§ 3.3.3 Krytyka teorii realistycznych

Hare podał jednak również inne racje za odrzuceniem mocnego kognitywizmu (realizmu moralnego). W eseju *Ontology in Ethics*, analizując cztery najbardziej typowe schematy językowe, w których uprawia się metaetykę (ontologiczny, psychologiczny, epistemologiczny i pojęciowy), stwierdził, iż „ontologiczny sposób stawiania tych kwestii jest dziś najbardziej popularny i zarazem najbardziej prymitywny”⁶².

⁶² R. M. Hare, *Ontology...*, dz. cyt., s. 83.

Hare nie interesował się metafizyką. Był w tej kwestii kwietystą: wychodził z założenia, że rozważania metafizyczne nie prowadzą do żadnych istotnych wniosków. Jego zdaniem, ważniejsze od sporów na temat sposobu istnienia wartości jest znalezienie metody zapewnienia obiektywności ocen oraz racjonalności dyskursu etycznego, gdyż etyczny racjonalizm nie wymaga etycznego realizmu.⁶³ Należy odrzucić metafizykę moralności, ponieważ, po pierwsze, gdy tylko zaczniemy zastanawiać się nad istnieniem wartości, pojawia się bardzo skomplikowane pytanie, co znaczy „istnieć”. Słowo to ma różne znaczenia, które często są ze sobą mylone. W pewnym sensie wartości moralne istnieją, lecz nie istnieją one tak jak np. stoły⁶⁴, gdyż istnienie wartości jest od nas zależne. Po drugie, jeśli założenia realistyczne przyczyniają się do rozwoju nauki (np. założenie istnienia kwarków w fizyce spowodowało postęp), to w dziedzinie etyki zalety takiego podejścia są złudne. Zakładanie realizmu nie wpływa bowiem w żaden pozytywny sposób na nasze zdolności poznania prawdziwości zasad moralnych czy na odpowiedź na konkretne pytanie: „Co powinienem w tej sytuacji uczynić?”. Tak jak prawideł metody naukowej nie można poznać wyłącznie na podstawie zdań empirycznych lub poprzez ustanowienie poznawczego kontaktu z bytami naukowymi, tak też o prawdziwości zasad moralnych nie dowiemy się niczego, gdy założymy istnienie moralnych własności. Dla etyka ontologia jest więc dziedziną nieistotną⁶⁵. Po trzecie, „gdy rozważamy tzw. ontologiczne zagadnienia etyki, poruszamy te same kwestie, które równie dobrze mogą być postawione za pomocą języka analizy pojęciowej czy logicznej”⁶⁶. Hare był przekonany, że w dziedzinie etyki wszystkie problemy ontologiczne w swoim głębszym wymiarze są problemami pojęciowymi. Filozofowie, zamiast uprawiać metafizykę, powinni zająć się językiem. To, co realisci moralni nazywają koniecznymi prawami metafizycznymi jest niczym innym jak substancjalnymi zasadami moralnym lub deontycznymi modalnościami⁶⁷. Zatem jeśli nie można jednoznacznie rozstrzygnąć, czy wartości etyczne istnieją, nie mamy bowiem

⁶³ Por. tamże, s. 92-93.

⁶⁴ Zob. tamże, s. 86.

⁶⁵ Zob. tenże, *Myślenie...*, dz. cyt., s. 146.

⁶⁶ Tenże, *Ontology...*, dz. cyt., s. 85.

⁶⁷ Zob. tenże, *A New Kind...*, dz. cyt., s. 85-86.

do nich poznawczego dostępu, a postulowanie ich istnienia nie przyczynia się do rozwiązania jakichkolwiek problemów praktycznych, powinniśmy budować teorie, które nie wymagają uwzględnienia ich istnienia. Hare skłaniał się więc ku antyrealizmowi.

Najcięższym zarzutem, jaki postawił teoriom realistycznym, było wskazanie na ich niepraktyczność⁶⁸. Ponieważ dziedziną zainteresowania realistów jest głównie ontologia, według niego nie potrafią oni przedstawić żadnej metody rozwiązywania ludzkich problemów moralnych, która nie byłaby oparta na spornych opiniach zależnych od nabytych przekonań czy intuicji. Teorie takie nie nadają się do rozwikłania np. konfliktów międzykulturowych. Gdy zaakceptujemy oczywisty fakt, iż istnieje rozbieżność poglądów moralnych, takie teorie nie mają narzędzi, aby się jej pozbyć⁶⁹. Co więcej, zdaniem Hare'a to deskryptywny realizm konserwuje taki stan rzeczy. W jaki sposób bowiem realista może przekonać tych, którzy nie podzielają jego opinii, że np. obrzezanie kobiet jest złe? Jeśli dokonywanie adekwatnego osądu moralnego polega na ustanowieniu kontaktu poznawczego z niezależnie ukonstytuowanym stanem rzeczy, to jakie są kryteria rozstrzygnięcia, czy ów kontakt poznawczy został adekwatnie ustanowiony? Ponieważ jest jasne, że ludzie nie mają żadnego nieomylnego zmysłu moralnego, intuicji, która za każdym razem adekwatnie korespondowałaby z własnościami moralnymi⁷⁰, musielibyśmy zawsze pozostawać w sferze subiektywnych lub co najwyżej społecznie usankcjonowanych przekonań. Podczas konfliktów moralnych każda ze stron może mieć jednak odmienne, mocne przekonanie, że to ona ma rację. Jeśli realisci nie przedstawiają jakiegos sposobu rozstrzygnięcia, które z dwu przeciwstawnych przekonań jest adekwatne, sposobu nieopartego jedynie na etnocentrycznych poczuciach i uprzedzeniach, często nazywanych „moralnymi truizmami” (*moral platitudes*) lub „naszymi żywotnymi przekonaniem”, to realizm nie będzie się wiele różnił od subiektywistycznej i relatywistycznej koncepcji emotywistów⁷¹. Zatem intuicjonizm⁷², który wśród

⁶⁸ Zob. tenże, *A 'Reductio...*, dz. cyt., s. 119.

⁶⁹ Zob. tenże, *A New Kind...*, dz. cyt., s. 82.

⁷⁰ Zob. tenże, *Myślenie...*, dz. cyt., s. 43-45, 56, 167.

⁷¹ Por. tenże, *How to Decide Moral Questions Rationally*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 104.

⁷² Intuicjonistami Hare nazywa wszystkich tych filozofów, którzy odwołują się do intuicji moralnych lub powszechnie akceptowanych przekonań moralnych

realistów jest najczęściej akceptowanym poglądem epistemologicznym dotyczącym sposobu poznawania własności moralnych, pociąga za sobą subiektywizm, co znów, inną drogą, prowadzi do relatywizmu⁷³. Jeśli natomiast wykluczmy intuicjonizm z grona dobrych metod zdobywania wiedzy, to, zdaniem Hare'a, realicie nie pozostanie już wiele sposobów na rozwiązanie konfliktów moralnych. Nawet jeśli jakaś społeczność posiada prawdziwe przekonania moralne, to przedstawiciel realizmu nie może mieć dobrych narzędzi pozwalających wykazać, iż te właśnie standardy oceniania są zgodne z rzeczywistością⁷⁴.

Powyższy sposób podważenia kognitywizmu mocnego wydaje się zbyt pochopny. Jeśli bowiem odrzucimy dyskusyjne założenia, że (1) teoria metafizyczna, zamiast być prawdziwa, powinna być praktyczna oraz że (2) każda semantyka deskrytywistyczna ma z konieczności negatywny wpływ na proces ustalania obiektywnych praw moralnych, możemy spróbować wyobrazić sobie systemy deskrytywistyczne lub realistyczne, które nie miałyby krytkowanych przez Hare'a konsekwencji epistemologicznych. Gdy zauważymy, że nie ma prostej relacji między deskrytywizmem a relatywizmem, można bowiem być realistą i semantycznym eksternalistą⁷⁵ (np. realizm z Cornell), oraz że nie ma prostej relacji pomiędzy realizmem a poznawczą niemocą, gdyż zamiast być intuicjonistą można wierzyć w naturalizm, realizm i w metodologiczny scjentyzm, okaże się iż,

jako do źródła wiedzy na temat zasad postępowania. Do zwolenników takiej teorii Hare zalicza J. Rawlsa, R. Nozicka, R. Dworkina, B. Williamsa, S. Hampshire'a czy H. Putnama. (Zob. tenże, *How to Decide...*, dz. cyt., s. 101.) W podręcznikach metaetyki jako intuicjonistów klasyfikuje się natomiast takich filozofów, jak R. Audi, C. D. Broad, J. Butler, E. F. Carrit, S. Clarke, R. Cumberland, J. Dancy, A. C. Ewing, F. Hutcheson, O. A. Johnson, G. E. Moore, R. Price, H. A. Prichard, H. Rashdall, Th. Reid, W. D. Ross, H. Shaftesbury. „Wszyscy intuicjoniści dzielą pogląd, iż jakakolwiek by mogła być prawda w moralności, musi ona być odkryta na drodze bezpośredniej intuicji czy wglądu moralnego” (M. Fritzhand, *Główne...*, dz. cyt., s. 93).

⁷³ Zob. R. M. Hare, *Sorting...*, dz. cyt., s. 90.

⁷⁴ Tamże, s. 138.

⁷⁵ **Semantyczny eksternalizm** to pogląd, według którego znaczenie wielu słów (treść wielu pojęć lub myśli) przynajmniej częściowo jest zależne od czynników zewnętrznych wobec podmiotu, takich jak środowisko fizyczne, środowisko społeczne czy kauzalne związki ze światem. Zob. R. Ziemińska, *Eksternalizm we współczesnej epistemologii*, Rozprawy i Studia US, t. 407, Szczecin 2002, s. 241.

powyższe argumenty przeciw deskrytywizmowi nie są do końca przekonujące.

§ 3.3.4 Krytyka scjentyzmu – Hare'a zmodyfikowany test otwartego pytania

Odrzuciwszy intuicję jako źródło poznania faktów moralnych, Hare w dziedzinie etyki nie zaakceptował również scjentyzmu. Przez scjentyzm rozumiemy pogląd, który głosi, iż na drodze badań naukowych można uzyskać odpowiedź na pytania moralne. Uważał on, iż nauka (psychologia, socjologia, lingwistyka itd.) nie może być adekwatnym źródłem poznania praw etycznych. W *The Language of Morals*⁷⁶ przedstawił – jak nazywa go wielu autorów – **Hare'a zmodyfikowany test otwartego pytania**⁷⁷. W jego oczach stanowił on główny argument uzasadniający rozdział między przekonaniem dotyczącym faktów a sądami moralnymi. Test przeprowadza się w następujący sposób. Załóżmy, że zdanie „To jest dobry obraz” możemy wydedukować z jakiegoś zbioru sądów opisowych; może mieć on formę pojedynczego opisu, koniunkcji lub dysjunkcji alternatywnych opisów. Nazwijmy ową deskrypcję C. Jako że według Hare'a dedukcja może mieć wyłącznie charakter analityczny, zdanie „To jest dobry obraz” musi wynikać ze zdania „To jest obraz C” na mocy znaczenia terminów w nim użytych. Zdanie „To jest dobry obraz” znaczy więc: „To jest obraz, który jest C”. Jednak gdyby tak było, nie moglibyśmy uchwycić sensowności wypowiedzi takiej jak „To jest dobry obraz, ponieważ jest C”, gdyż po podstawieniu „doby obraz”/„obraz C” otrzymalibyśmy „To jest obraz C, ponieważ jest C”. Jest oczywiste, że nie to chcielibyśmy powiedzieć, definicja powyższa nie oddaje więc znaczenia wypowiedzi „To jest dobry obraz”, ponieważ zdanie „To jest

⁷⁶ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 84-85.

⁷⁷ Opis i krytykę Hare'a wersji testu otwartego pytania znaleźć można w: R. Hancock, *The Refutation of Naturalis in Moore and Hare*, „The Journal of Philosophy” (10) 1960, s. 326-334; J. F. Lange, R. M. Hare's *Reformulation of the Open Question*, „Mind” (298) 1966, s. 244-247; S. Stojanovic, *Hare's Argument Against Ethical Naturalism*, „Mind” (286) 1963, s. 264-267; R. Hancock, *A Note on Hare's The Language of Morals*, „The Philosophical Quarterly” (50) 1963, s. 56-63; L. W. Sumner, *Hare's Arguments Against Ethical Naturalism*, „The Journal of Philosophy” (23) 1967, s. 779-791; Ch. Daniels, *Hare on the Meaning of „Good”*, „Mind” (313) 1970, s. 139-141; A. Miller, *An Introduction...*, dz. cyt., s. 21-22.

obraz C" nie jest jej analitycznym ekwiwalentem. Wypowiadając „To jest dobry obraz” – twierdził dalej Hare – tak naprawdę chcieliśmy zarekomendować obraz, który jest C. To natomiast, że polecamy jego obejrzenie, nie jest równoznaczne temu, że jest on C, choć C może być racją tego, że go polecamy. Powyższy *zmodyfikowany test otwartego pytania Hare'a* za A. Millerem zrekonstruować można w kilku bardziej ogólnych punktach⁷⁸:

Założmy, że:

(1) Predykat „dobry” jest analitycznym odpowiednikiem naturalnego predykatu „N”.

(2) Istnieje wewnętrzne (pojęciowe) połączenie między wydawaniem sądu moralnego a posiadaniem motywacji do działania.

Zatem na mocy definicji analityczności wiemy, że:

(3) Znaczenie tezy „x jest N” jest identyczne ze znaczeniem tezy „x jest dobry”.

Jeśli tak, to:

(4) Ktoś, kto poważnie zadaje pytanie „Czy osoba O może nie być motywowana do wykonania działania, o którym stwierdza «x jest N»?”, popełniłby błąd pojęciowy.

Jednakże nie wydaje się, aby tak było (4), gdyż:

(5) Kompetentni i rozumni użytkownicy języka polskiego są przekonani, iż mogą wyobrazić sobie w pełni świadomą i psychicznie zdrową osobę, która twierdząc, że „x jest N”, nie jest motywowana lub nie posiada racji do działania zgodnego z tym sądem.

(6) To, że kompetentni użytkownicy mieliby przekonanie opisane w punkcie (5), wskazuje, iż nie istnieje pojęciowe połączenie między sądem „x jest N” i motywacją do wykonania zgodnego z nim działania.

Zatem:

(7) Pytanie: „Czy osoba O może nie być motywowana do wykonania działania, o którym stwierdza «x jest N»?” jest pytaniem *otwartym* – znaczącym i nie jawi się nam jako pojęciowe pomieszanie.

Zatem:

(8) Jeśli nie istnieje lepsze wyjaśnienie przekonania opisanego w punkcie (5), musimy stwierdzić, że nie istnieje pojęciowe połą-

⁷⁸ Por. tamże. Por. R. M. Hare, *The Language...*, dz. cyt., s. 84-85. Znacznie wcześniej podobną argumentację przedstawił D. Hume; zob. D. Hume, *Traktat...*, dz. cyt., s. 244, 250, oraz M. Rutkowski, *Rola...*, dz. cyt., s. 86-87.

czenie między sądem „x jest N” i motywacją do wykonania działania z nim zgodnego.

Zatem na mocy punktu (3):

(9) Sąd „x jest N” nie jest sądem moralnym.

Zatem:

(10) „N” nie jest analitycznym odpowiednikiem „dobry” („dobry” i „N” nie są synonimami).

Zmodyfikowana wersja *testu otwartego pytania* różni się znacznie od swego oryginału⁷⁹. Można zauważyć, iż zmianie uległ nie tylko sam argument, ale również dziedzina jego zastosowania. Uderza on bowiem nie tylko w naturalizm definicyjny, lecz w każdą formę deskrytywizmu (również w nonnaturalizm i naturalizm semantyczny). Zdaniem Hare'a, deskrytywista nie może wygłaszać sądów normatywnych, ponieważ odpowiedź na pytanie „Co powinienem zrobić” przynajmniej w części musi być zalecaniem – musi wpływać na nasze wybory⁸⁰. Z tego powodu „żaden sąd moralny nie może być czystym stwierdzeniem faktu”⁸¹. Jeśli uważamy, że sądy moralne są jedynie stwierdzeniami, uniemożliwiamy im wypełnienie ich głównej funkcji – kierowanie naszymi działaniami⁸². U podłoża tego argumentu leży przeświadczenie, iż nauka i etyka to dwie różne dziedziny. Pierwsza z nich jest działalnością deskrytywną, której głównym celem jest obiektywne i prawdziwe przedstawienie rzeczywistości. Jej sądy są sprawdzalne empirycznie. Etyka natomiast jest działalnością normatywną, której wynikiem powinny być zalecenia, one bowiem mogą motywować do działania. Jej celem jest podawanie odpowiedzi na pytanie „jak żyć?” i „co mamy czynić?”, nie zaś opis świata, choćby świata moralności. Etyka nie jest dziedziną odkrycia i opisu, ale rozumowania, dialogu i norm.

Jednak powyższe założenie, choć jest akceptowane przez nondeskrytywistów, nie jest tezą oczywistą. Po pierwsze, zakłada prawdziwość internalizmu motywacyjnego, choć niektórzy filozofowie wytaczają przeciwko niemu przekonujące argumenty⁸³. Po drugie,

⁷⁹ Zob. § 2.1.

⁸⁰ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 29.

⁸¹ Tamże.

⁸² Zob. tamże, s. 44, 46.

⁸³ Zob. § 3.17.

nie udowadnia niemożności motywowania przez sądy opisowe, a jedynie, w związku z trudnościami tego projektu, stawia nieprzychylną dla deskryptywizmu hipotezę. Choć podanego przez Hare'a zdania: „To jest dobry obraz” jak dotąd nie udaje się w satysfakcjonujący sposób sprowadzić do opisu, wyciąganie stąd wniosku, że jest ono preskrypcją i jego znaczenie nie jest wyznaczone tylko i wyłącznie przez warunki prawdziwości, wydaje się tezę *ad hoc*. Są metaetycy, którzy bronią relacji pomiędzy motywacją i specyficznymi sądami opisowymi⁸⁴, a dyskusja motywacyjnych internalistów z eksternalistami daleka jest od powszechnie akceptowanych rozstrzygnięć⁸⁵.

§ 3.4 Kryteria oceny teorii metaetycznych

W artykule *Fundationalism and Coherentism*⁸⁶ Hare rozróżnił dwa rodzaje fundacjonalizmu: linearny – nazywany także fundacjonalizmem kartezjańskim oraz hipotetyczno-dedukcyjny – wzorowany na podejściu K. R. Poppera⁸⁷ i W. O. Quine'a⁸⁸, przy czym w *Mysleniu moralnym* przedstawił siebie jako metodologicznego popperystę, a nie kartezjanistę. Jego teorię traktować trzeba jako zbiór hipotez, które okazują się trafne, jeśli przechodzą odpowiednie testy – nie opiera się ona na jasnych i wyraźnych przesłankach. Podejście to stosował do swoich tez metaetycznych⁸⁹ oraz konkretnych reguł etycznych wyprowadzanych na płaszczyźnie etyki normatywne⁹⁰. To, że sądy etyczne są zuniwersalizowanymi preskrypcjami, nie jest więc definicją sprawozdawczą, lecz hipotezą lingwistyczną⁹¹.

Dobra hipoteza według Poppera powinna jednak spełniać szereg warunków, np. nie może być idealna. Teoria jest idealna, jeśli nie można jej sfalsyfikować⁹². Choć system Hare'a również musi bronić

⁸⁴ Wśród współczesnych internalistów akceptujących deskryptywizm wymienić można np. J. McDowella, J. L. Mackiego, T. Nagela, M. Smitha, D. Wigginsa.

⁸⁵ Zob. M. Rutkowski, *Obowiązek moralny a motywacja*, Wydawnictwo Naukowe Uniwersytetu Szczecińskiego, Szczecin 2001.

⁸⁶ R. M. Hare, *Fundationalism...*, dz. cyt., s. 115-125.

⁸⁷ Zob. K. R. Popper, *Logika odkrycia naukowego*, tłum. U. Niklas, PWN, Warszawa 1977.

⁸⁸ W. Quine, *Dwa dogmaty empiryzmu*, w: tenże, *Z punktu widzenia logiki*, tłum. B. Stanosz, PWN, Warszawa 1969.

⁸⁹ R. M. Hare, *Myslenie...*, dz. cyt., s. 26.

⁹⁰ Zob. tenże, *Fundationalism...*, dz. cyt., s. 117.

⁹¹ Zob. tenże, *Myslenie...*, dz. cyt., s. 26-27, 87.

się przed takim zarzutem, komentatorzy wskazują, iż nie udaje mu się uniknąć powyższego problemu. Thomas Nagel stwierdził, iż teoria Hare'a „zbyt dobrze sobie radzi”⁹³. Alasdair MacIntyre w *Krótkiej historii etyki* również stawia mu podobny zarzut. Prezentując na kartach tej książki paradygmat etyki, którego zwolennikiem był Hare, stwierdził, że w owym podejściu:

Najpierw przedmiotem uwagi staje się pewna niesystematycznie wybrana klasa pojęć moralnych i sądów. Z tak przeprowadzonego studium wyprowadza się wnioski, że dyskurs specyficznym moralny ma pewne określone cechy. Gdy ktoś przytoczy kontrprzykłady, odrzuca się je jako nieistotne, ponieważ nie są przykładami dyskursu moralnego; z kolei faktu, że nie mają one moralnego charakteru, dowodzi się, wskazując, że nie posiadają koniecznych cech przysługujących elementom dyskursu moralnego. Z tego błędnego koła wybawić nas może tylko adekwatny historyczny pogląd na różnorodność rodzajów dyskursu moralnego i wartościującego.⁹⁴ [...] Istotna w tym miejscu jest pewna niejasność w całym przedsięwzięciu Hare'a, którą wskazała Mary Warnock⁹⁵. Czy Hare, definiując wartościowanie i preskrypcję, czyni to w taki sposób, aby ocalić swoją tezę przed możliwymi kontrprzykładami? [...] Jeśli ustala arbitralnie, że ocena i zakaz (preskrypcja) są tym, co on uznaje za ocenę lub zakaz, dlaczego mielibyśmy zaakceptować jego ustalenie? [...] W istocie wydaje się on polegać na niemal intuicyjnym rozumieniu tego, co należy włączyć, a co wykluczyć z dziedziny wyrażenia wartościujących⁹⁶.

Zatem rozstrzygnąć należy pytania: czym dla systemu Hare'a mogłyby być potencjalny falsyfikatory⁹⁷? Co ów system mogłoby wery-

⁹² „Naturalnie tylko wówczas traktuję pewien system jako empiryczny lub naukowy, gdy poddaje się on sprawdzeniu w doświadczeniu. Z rozważań tych wynika, że kryterium demarkacji należy przyjąć nie weryfikowalność, lecz falsyfikowalność systemu. Innymi słowy, nie wymagam, by jakiś system naukowy można było wybrać raz na zawsze w sensie pozytywnym, wymagam natomiast, by miał on taką formę logiczną, aby testy empiryczne pozwalały na decyzję w sensie negatywnym: musi być możliwe obalenie empirycznego systemu naukowego przez doświadczenie”; K. R. Popper, *Logika...*, dz. cyt., s. 39-40.

⁹³ T. Nagel, *The Foundations of Impartiality*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 101. Krytyka T. Nagela zanurzona jest w nieco innym kontekście. Nagel w ten sposób próbuje krytykować Hare'a za to, iż nie potrafi on wyjaśnić istnienia moralnej niezgody. Istnienie więc moralnej niezgody według Nagela byłoby czymś, co falsyfikuje teorię Hare'a.

⁹⁴ A. MacIntyre, *Krótko...*, dz. cyt., s. 32.

⁹⁵ M. Warnock, *Ethics Since 1900*, Oxford University Press, Oxford 1960, s. 128-129.

⁹⁶ A. MacIntyre, *Dziedzictwo...*, dz. cyt., s. 327.

⁹⁷ Przez potencjalne falsyfikatory rozumiemy takie logicznie możliwe zdania, które jeśli byłyby prawdziwe, mogłyby sfalsyfikować teorię Hare'a. Zob. K. R. Popper, *Logika...*, dz. cyt., s. 74-88.

fikować? Dla Poppera o wartości teorii świadczy między innymi możliwość jej sfalsyfikowania, jej moc eksplanacyjna i przewidywalność. Jak mają się te kryteria do uniwersalnego preskryptyzmu?

W *Myśleniu moralnym* Hare, podsumowując wywód na temat swojego podejścia metodologicznego, pisał:

Krótko mówiąc, wolno nam tymczasem zachować te hipotezy na temat znaczeń słów, które – jeśli to prawda i jeśli nie mylimy się co do tego, jacy są ludzie – odpowiadałyby za to, jakie poglądy moralne faktycznie wyznają⁹⁸.

Zatem wydaje się, iż potencjalnymi weryfikatorami systemu etycznego powinny być istniejące w społeczeństwie zasady moralne; teorię weryfikujemy poprzez wyprowadzanie z niej sądów moralnych i porównywanie ich z faktycznymi ocenami funkcjonującymi w społeczeństwie; jej moc eksplanacyjną moglibyśmy mierzyć tym, jak wyjaśnia ona istnienie i kształt zastanych w społeczeństwie przekonań⁹⁹; jej moc przewidywalną rozumieć jako ocenę adekwatności postulowanych przez teorię sądów etycznych odnoszących się do sytuacji hipotetycznych.

Czy istnieje jednak jakiś odpowiednik Popperowskich potencjalnych falsyfikatorów, czegoś, co mogłoby nas przekonać, że teoria nie jest trafna? Pytanie to jest ważne, gdyż jedynie ich istnienie zabezpieczy ją może przed idealnością. Do miana falsyfikatorów mogłyby pretendować podzielane w społeczeństwie intuicje moralne, które byłyby sprzeczne z zasadami wyprowadzonymi z etyki normatywnej Hare'a. Twierdził on jednak, że nie można za pomocą intuicji moralnych i obowiązujących nakazów podważać zasad wyprowadzonych przez myślenie krytyczne¹⁰⁰, gdyż to ono jest nadrzędne i ono ustala zasady domniemane¹⁰¹. Rolę falsyfikatorów mogą pełnić jedynie dane lingwistyczne. Hare bowiem, prezentując uniwersalny preskryptyzm, starał się oddać faktyczny sposób, używania pojęć. Według niego,

jeśli próbujemy wyjaśnić znaczenie słów moralnych zgodnie z tym, jak są używane, to sposób w jaki są używane, jest priorytetowy dla poprawności lub niepoprawności naszego wyjaśnienia. Jeśli twierdzimy, że słowa są używane w sposób, w który nie są używane, wtedy po prostu jesteśmy w błędzie¹⁰².

⁹⁸ R. M. Hare, *Myślenie...*, dz. cyt., s. 26.

⁹⁹ Zob. tamże, s. 25-26.

¹⁰⁰ Pojęcia myślenia krytycznego (zasad krytycznych) oraz myślenia intuicyjnego (zasad domniemanych) wyjaśnione zostały w §3.3 oraz §6.1.

¹⁰¹ Zob. tamże, s. 63, 172.

¹⁰² Tenże, *The Argument from Received Opinion*, w: tenże, *Essays on Philosophical...*, dz. cyt., s. 122.

Oceniając teorie metaetyczne, Hare kierował się konkretnym zbiorem wymagań, które powinny być spełnione, aby dany system można było nazwać dobrym. Wpływał on, po pierwsze, z przekonania, iż metaetyka powinna umożliwiać etyce normatywnej spełnianie swojej funkcji praktycznej, oraz, po drugie, z obserwacji życia moralnego: zadawanych pytań, oczekiwań w stosunku do odpowiedzi, sposobu użycia zwrotów etycznych, właściwości zdań i sądów moralnych, roli wypowiedzi moralnych w życiu poszczególnych ludzi, charakteru sporów, sposobu uzasadniania maksym, wagi przypisywanej moralności. Podając zaprezentowaną w § 3.2 klasyfikację stanowisk metaetycznych, próbował zbadać, czy realizują one „listę wymagań, którą adekwatna teoria etyczna musi spełnić”¹⁰³. Zestawienie owych kryteriów wygląda następująco¹⁰⁴ (plusami zaznaczono zostały teorie, które spełniają dane kryterium, minusami zaś te, które go nie spełniają:

	Obiektywistyczny naturalizm	Subiektywistyczny naturalizm	Intuicjonizm	Emotywizm
1. Neutralność	-	+	+	+
2. Praktyczność	-	-	-	+
3. Niezgodność	+	-	+	+
4. Logiczność	+	+	+	-
5. Argumentacyjność	-	-	-	-
6. Mediesyjność	-	-	-	-

Tabela 2. Kryteria oceny teorii metaetycznych

Wymagania przedstawione w tabeli Hare objaśnił w następujący sposób¹⁰⁵:

Neutralność (*neutrality*) – zakaz wprowadzania materialnych opinii moralnych do teorii pod pozorem definicji terminów moralnych. Warunek ten wydaje się być słabszą wersją założenia, że teoria metaetyczna musi być formalna. Intencją Hare'a było rozgraniczenie metaetyki od etyki. Rozwiązania metaetyczne powinny móc być podzielane przez każdą ze stron dyskusji etycznej, dlatego też muszą być one wolne od jakichkolwiek materialnych założeń etycznych.

¹⁰³ Tenże, *Sorting...*, dz. cyt., s. 46.

¹⁰⁴ Tamże, s. 42.

¹⁰⁵ Tamże, s. 118-122.

Praktyczność (*practicality*) – uznanie sądu moralnego ma być zarazem racją do działania. Kryterium to w konsekwencji oznacza, iż dobra teoria powinna być zgodna z internalizmem motywacyjnym.

Niezgodność (*incompatibility*) – dobra teoria etyczna powinna traktować spory moralne jako autentyczne spory normatywne, a nie dyskusje na temat znaczeń. Przynajmniej niektóre wygłaszane przez ludzi sądy moralne muszą być ze sobą niezgodne. Warunek ten zabezpiecza teorię przed relatywizmem normatywnym i argumentem *reductio ad absurdum*.

Logiczność (*logicality*) – pomiędzy sędziami moralnymi powinny zachodzić relacje logiczne (np. sprzeczność, wynikanie, rozłączność itd.). Kryterium to Hare wprowadził, aby zabezpieczyć racjonalizm dyskursu etycznego.

Argumentacyjność (*arguability*) – teoria moralna powinna pozwalać rozwiązać przynajmniej niektóre praktyczne problemy moralne za pomocą argumentów i dyskusji.

Mediacyjność (*conciliation*) – powinna także za pomocą argumentów spełniać swoją funkcję społeczną, jaką jest godzenie sprzecznych interesów i stanowisk. Mediacyjność jest wyrazem optymistycznego podejścia Hare'a do problemów etyki.

Powyższe kryteria można sprowadzić do ogólnego zalecenia, które głosi, że teoria etyczna powinna umożliwiać rozwiązywanie praktycznych ludzkich problemów etycznych na drodze racjonalnych argumentów umożliwiających dyskusję, której wynikiem powinny być reguły moralne, będące motywem naszego działania. Warunki te korespondują z wysuwaniem przez Hare'a głównymi zarzutami skierowanymi przeciwko konkurencyjnym teoriom metaetycznym: z argumentacyjnością, mediacyjnością i niezgodnością – argument *reductio ad absurdum*; z logicznością – najważniejszy zarzut przeciwko emotywizmowi (zarzut irracjonalizmu); z praktycznością – *Hare'a zmodyfikowany test otwartego pytania*; z neutralnością – krytyka intuicjonizmu. To również z powodu powyższych kryteriów Hare skłaniał się ku uniwersalnemu preskryptywizmowi. Twierdził bowiem, że jedynie racjonalistyczny nondeskryptywizm może przejść wszystkie testy¹⁰⁶. Nazwiemy je **warunkami pragmatycznymi**, każda bowiem teoria, która je spełni, umożliwi realizację funkcji języka etycznego – potrzeb, które stały się przyczyną jego powstania.

¹⁰⁶ Zob. tamże, s. 118, 136-137.

Jednak powyższe kryteria nie są jedynymi, jakie wysuwali metaetycy. Również deskryptywiści, krytykując nonkognitywizm, przedstawili szereg własnych propozycji¹⁰⁷. Aby rzetelnie ocenić teorie metaetyczne, powinniśmy rozważyć je z jak najszerzej perspektywy, dlatego też weźmiemy również pod uwagę **kryterium fenomenologii moralnej** oraz **koherencji**¹⁰⁸. Pierwsze z nich stwierdza, że „teoria metaetyczna powinna, jeśli to możliwe, wyjaśnić i obronić tak wiele głęboko zakorzenionych aspektów myślenia i dyskursu moralnego – fenomenów moralności – jak to tylko możliwe.”¹⁰⁹ Drugie natomiast głosi, iż „teoria metaetyczna powinna pasować do budzących zaufanie założeń filozofii oraz innych dziedzin dociekań, włączając w to naukę”¹¹⁰. **Kryterium fenomenologii (KF)**, zdaniem M. Timmons'a i T. Horgana, wymaga od teorii metaetycznej wyjaśnienia następujących fenomenów moralności¹¹¹:

KF1 Sądy moralne standardowo są wypowiedzane w trybie oznajmującym, mają formę stwierdzeń i wydają się wyrażać autentyczne przekonania.

KF2 Sądy moralne mogą być prawdziwe lub fałszywe i faktycznie wiele z nich jest prawdziwych.

KF3 Mogą istnieć głębokie moralne różnice poglądów – włączając w to różnice pomiędzy fundamentalnymi moralnymi przekonaniami i zasadami.

KF4 Sądy moralne kierują naszym zachowaniem; charakterystyczne dla nich jest to, że motywują nas do działania.

KF5 Sądy moralne bazują na autorytecie rozumu. Oceny wydajemy opierając się na racjach.

KF6 Sądy moralne dotyczą ważnych i podstawowych zagadnień naszego życia, podczas gdy, np. sądy estetyczne czy reguły etykiety są, generalnie rzecz ujmując, kwestią mniejszej wagi.

Niektóre z powyższych fenomenów odpowiadają kryteriom zaproponowanym przez Hare'a: KF3 – *niezgodności*, KF4 – *praktyczności*, KF5 – *argumentacyjności*; KF6 zawarte jest natomiast w jego

¹⁰⁷ Zob. R. B. Brandt, *Etyka...*, dz. cyt., s. 385-396.

¹⁰⁸ Warunki te znaleźć można w: M. Timmons, *Morality without Foundations...*, dz. cyt., s. 11-15, oraz M. Timmons, T. Horgan, *Morality without Moral Facts*, w: J. Dreier (ed.), *Contemporary Debates in Moral Theory*, Blackwell Publishers 2006, s. 220-238.

¹⁰⁹ Tamże.

¹¹⁰ Tamże.

¹¹¹ Por. tamże.

definicji sądów moralnych¹¹². Hare nie uważał za ważne wyróżnienie warunków KF1 i KF2. Stanowią one jednak powód wysuwania najpoważniejszych zarzutów wobec nondeskrytywizmu: KF1 związane jest bowiem z argumentem Fregego-Geacha¹¹³, natomiast KF2 z zarzutem niepodatności na prawdę sądów etycznych¹¹⁴. Dlatego też podanie odpowiedzi na pytanie, czy Hare potrafił wyjaśnić fenomeny KF1 i KF2, będzie jednym z celów niniejszej rozprawy.

Kryterium koherencji (KH) według Timmonsa prowadzi się do trzech warunków¹¹⁵. Choć Hare *explicite* ich nie wymienił, wydaje się, iż również traktował je jako ważne. Na ich podstawie krytykował bowiem m.in. realizm, powołując się na J. L. Mackiego *argument z dziwaczności* (wskazywanie na sprzeczność realizmu z KH1–KH3):

KH1 Ontologicznie pierwotne są wszystkie te byty (własności, fakty, itd.), które są przedmiotem nauki.

KH2 Każdy dyskurs pozanaukowy musi być dostosowany do naukowego (naturalistycznego) obrazu świata.

KH3 Wiedza i uzasadnienie związane z dyskursem pozanaukowym muszą być zgodne z naturalistycznym sposobem wyjaśniania, którego używamy, gdy zastanawiamy się nad poznawczym dostępem, wiedzą czy uzasadnieniem przekonań na temat naturalnego świata fizycznego.

Wszystkie wymienione powyżej warunki (pragmatyczne, fenomenologiczne i koherencji) powinny stanowić kryteria oceny teorii Hare'a. Również rozwiązanie przedstawionych powyżej problemów emotywizmu będzie wskaźnikiem powodzenia jego przedsięwzięcia. Każdy system nondeskrytywistyczny powinien bowiem, oprócz spełniania powyższych kryteriów, charakteryzować się także następującymi cechami: opisać postawy, które są specyficzne dla języka moralnego; wyjaśnić, dlaczego wydając oceny, nie musimy posiadać w stosunku do nich żadnych postaw emocjonalnych; opisać relację występującą pomiędzy elementem deskryptywnym a pozadeskryptywnym w znaczeniu sądów; udowodnić, że zaproponowane nowe definicje wyrażań języka moralnego nie są definicjami projektującymi, ale sprawozdawczymi (udowodnić np., że każda uniwersalna preskrypcja jest sądem moralnym).

¹¹² Chodzi o nadrzędność. Zob. § 4.14.

¹¹³ Zob. § 4.7.

¹¹⁴ Zob. § 5.3.

¹¹⁵ Zob. Timmons, *Morality Without Foundations...*, dz. cyt., s. 13-14.

§ 4 Preskrytywizm

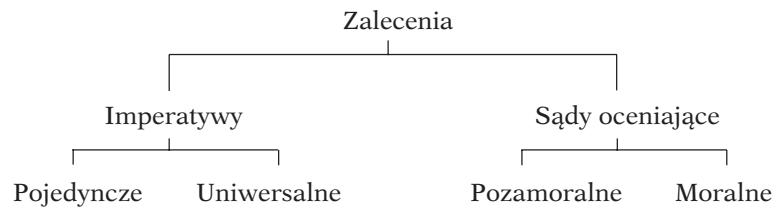
Najbardziej znaną tezę Richarda Hare'a jest stwierdzenie, że sądy moralne są zaleceniami. W niniejszym rozdziale przyjrzymy się założeniom, które leżą u jej źródeł, związanym z nią konsekwencjom oraz problemom, które stoją przed preskrytywistą. Ponieważ, zdaniem Hare'a, budowanie adekwatnej teorii normatywnej musi zostać poprzedzone przez zrozumienie jej języka, prześledzimy jego odpowiedzi na pytania: jakie znaczenie mają terminy moralne i jakimi cechami charakteryzują się sądy i jakimi cechami charakteryzują się relacje występujące pomiędzy pojęciami a sądami moralnymi. Uczynimy to, porównując sądy moralne z imperatywami oraz opisując je w języku teorii czynności mowy zarówno Austina–Searle'a, jak i samego Hare'a. Następnie zapoznamy się z najpoważniejszymi zarzutami skierowanymi pod adresem preskrytywizmu oraz odpowiedziami, które można na nie udzielić. Przedstawimy również Hare'a stanowisko w kwestii problemu przejścia od „jest” do „powinien” oraz relacji powinności do motywacji. Ostatnie zagadnienie zanalizujemy w odniesieniu do problemu słabości woli, amoralizmu oraz nadrzędności sądów moralnych.

§ 4.1 Preskrypcje

Nondeskrytywizm, a zatem również teorię Richarda Hare'a, zalicza się do **nurtu rewizjonistycznego** – tzw. *underlying-semantics revisionism*¹. Zgodnie z tym poglądem właściwa semantyka dyskursu

¹ Zob. A. Miller, *An Introduction...*, dz. cyt., s. 183.

moralnego różni się od semantyki, która jest sugerowana przez strukturę powierzchniową składni i gramatyki. Choć wydaje się, iż język moralności ma takie same właściwości semantyczne jak język opisujący stany rzeczy, jest to jedynie epifenomen. Różnica objawia się przede wszystkim w znaczeniu sądów i pojęć moralnych: ich znaczenie oprócz składni i warunków prawdziwości jest zdeterminowane także czymś więcej². Ową dodatkową, a zarazem istotną cechą jest preskryptywność. **Preskrypcjami (zaleceniami)** Hare nazywa pewną specyficzną klasę czynności mowy, którą cechują szczególne własności. Preskrypcje najczęściej mają charakter: komend, pochwał, pouczeń, instrukcji, modlitw, rozkazów, rad, prośb, błagań, zaproszeń, przepisów, nakazów itd.³ W przeciwieństwie do wypowiedzi, których funkcją jest przekazywanie, że coś przedstawia się tak a tak, służą one do mówienia komuś, że coś ma zrobić⁴. Istnieje kilka rodzajów wypowiedzi, które można zaliczyć do preskrypcji⁵, co przedstawia poniższy schemat:



Rysunek 3. Rodzaje zaleceń

Hare w różnych okresach życia w odmienny sposób objaśniał preskryptywny, ewaluatywny aspekt języka moralnego. Na początku swojej działalności posługiwał się gramatyką ogólną, w ostatnich latach natomiast (np. w *Sorting Out Ethics*) stosował już trzy odmiennie schematy pojęciowe: autorską teorię czynności mowy⁶,

² R. M. Hare, *Sorting...*, dz. cyt., s. 41. Zob. § 3.2.

³ Zob. tenże, *Imperatives, Prescriptions and Their Logic*, w: tenże, *Objective...*, dz. cyt., s. 43.

⁴ Zob. tamże.

⁵ Por. tenże, *The Language...*, dz. cyt., s. 3.

⁶ Zob. § 4.8.

teorię czynności mowy w wersji zaproponowanej przez J. L. Austina i J. Searle'a⁷, oraz warunkowoprawdziwością teorię znaczenia⁸, której użył do klasyfikacji systemów metaetycznych⁹. Owe odrębne języki, mówiące zarazem o znaczeniu, gramatycznych trybach i czynnościach mowy, prowadzą do swoistych kłopotów.

§ 4.1.1 Analiza gramatyczna sądów moralnych

W *The Language of Morals* Hare przedstawił sądy moralne poprzez analogię do imperatywów, imperatywy należą bowiem do zbioru preskrypcji i dzielą z zasadami moralnymi ważne cechy logiczne. Używając języka gramatyki, chciał on wykazać podobieństwa i różnice występujące między nimi. Według niego każde zalecenie charakteryzuje się trybem imperatywnym. Wyrażenia, które nazywał imperatywami czy komendami, nie są jednak tym, co powszechnie określa się tą nazwą. Hare znaczenie owego terminu traktował szerzej. Wyróżnił sześć rodzajów sądów, które posiadają tryb imperatywny¹⁰. Przedstawiono je w poniższej tabeli 3:

Imperatywy	
A	Imperatywy pojedyncze: <i>Odłóż mi pierniki!</i>
A'	Imperatywy uniwersalne niepełne: <i>Nie wolno pić!</i>
ZTI	Imperatywy zmodyfikowane: <i>! (Aż zaniknął dźwięk)</i>
Sądy oceniające	
B	Zalecenia odnoszące się do danego rodzaju sytuacji: <i>Powinno się zawsze odładować pierniki, jeśli się odlicza, że się je odlicza.</i>
C	Zalecenia o charakterze zarówno A i B (odnoszą się bezpośrednio do sytuacji jednostkowej, presuponują jednak istnienie preskrypcji B): <i>Powinnoś odładować mi pierniki.</i>
D	Zalecenia o charakterze C, aleppow ocenił: <i>Powinnoś być odładować mi pierniki.</i>

Tabela 3. Rodzaje preskrypcji

⁷ J. Searle, *Czynności...*, dz. cyt.

⁸ Jak wspomnieliśmy powyżej, bardziej poprawne byłoby nazwanie jej justyfikacyjną teorią znaczenia.

⁹ Zob. § 3.2.

¹⁰ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 155-157.

Najbardziej typowymi przedstawicielami zaleceń są wypowiedzi, posiadające formę imperatywu typu A. Klasycznym przykładem „zwykłego” imperatywu A jest rozkaz wojskowy, np. *Naprzód marsz!* Aby przedstawić analogię pomiędzy sądami moralnymi a imperatywami, Hare skonstruował **zmodyfikowany tryb imperatywny (ZTI)**. Zaliczał do niego również takie czynności mowy, jak instrukcje, rady, prośby, zaproszenia, a nawet modlitwy. Co więcej, postulował, aby w tym trybie imperatywy mogły występować w liczbie pojedynczej i mnogiej, we wszystkich osobach oraz czasach¹¹. Zmodyfikowany tryb imperatywny różni się od imperatywów klasycznych (A i A') w dwóch zasadniczych punktach: standardowe imperatywy, w przeciwieństwie do zmodyfikowanego trybu imperatywnego, występują najczęściej jedynie w formie drugiej osoby oraz nie przybierają formy czasu przeszłego. Pojęcie zmodyfikowanego trybu imperatywnego było pierwszą próbą Hare'a przedstawienia struktury głębokiej sądów etycznych w aspekcie ich preskryptywności.

Zapiszmy najbardziej charakterystyczne przykłady imperatywów w formie zmodyfikowanej. Niech znak „!” będzie symbolem takich wyrażań, których szczere użycie implikuje zanurzenie danej wypowiedzi w strukturze logicznej ZTI¹²:

Liczba pojedyncza	Liczba mnoga
! Ja zamykam drzwi. (<i>Niech ja zamykam drzwi!</i>) ! Ty zamykasz drzwi (<i>Zamykaj drzwi!</i>) ! On/Ona/Ono zamyka drzwi (<i>Niech On/Ona/Ono zamyka drzwi!</i>)	! My zamykamy drzwi. (<i>Zamknijmy drzwi!</i>) ! Wy zamykacie drzwi (<i>Zamknijcie drzwi!</i>) ! Oni zamykają drzwi (<i>Niech Oni zamkną drzwi!</i>)
Czas przyszły	Czas przeszły
! Ja zamknę drzwi. (<i>Niech ja zamknę drzwi!</i>) ! Ty zamkniesz drzwi (<i>Zamknij drzwi!</i>) ! On/Ona/Ono zamknie drzwi (<i>Niech On/Ona/Ono zamknie drzwi!</i>)	! Ja zamknąłem drzwi. (<i>Ja powinienem był zamknąć drzwi!</i>) ! Ty zamknąłeś drzwi (<i>Ty powinienś był zamknąć drzwi!</i>) ! Oni zamknęli drzwi (<i>Oni powinni byli zamknąć drzwi!</i>)

Tabela 4. Przykłady zmodyfikowanego trybu imperatywnego

¹¹ Zob. tamże, s. 188-190.

¹² Hare nie używał ani symbolu „!” ani „□”, zamiast nich stosował natomiast

Hare nie podał ścisłej definicji ZTI; zamiast tego prezentował typowe przykłady jego użycia, próbując kontekstowo objaśnić to pojęcie. Operator „!” traktował jako funktor zdaniotwórczy od argumentów zdaniowych, za pomocą którego ze zdań deskryptywnych możemy tworzyć pewne rodzaje zaleceń. „!” pełni więc rolę modyfikatora gry językowej oraz wskaźnika preskryptywności wypowiedzi.

Słowo „powinno się” posiada znaczenie podobne do „!”. Jest ono również funktorem zdaniotwórczym od argumentów zdaniowych. Wprowadza jednak inne relacje niż „!”, ponieważ preskrypcje B, C i D posiadają dodatkowe własności formalne. Gdy np. zalecenie „!” może być jedynie uniwersalne niepełnie, „powinno się” (zapiszmy to pojęcie za pomocą symbolu „□”) musi być uniwersalne pełnie¹³ oraz mieć uzasadnienie¹⁴. Sposób użycia „□” przedstawił Hare w następująco:

Powinno się¹⁵ (□)

I. Znaczenie (B)

- i. „Wszystkie P powinny być Q” = „Wszystkie P są Q, proszę” = „□ Wszystkie P są Q”
- ii. „Wszyscy powinni zawsze oddawać pieniądze” = „Każdy oddaje pieniądze, proszę” = „□ Każdy oddaje pieniądze”

II. Znaczenie (C) i (D)

- i. „Powinieneś powiedzieć mu prawdę” = „Jeżeli nie powiesz mi prawdy, złamiesz podstawową (*general*) powinność – zasadę, pod którą się niniejszym podpisuję” = „□ powiesz mi prawdę”
- ii. „Powinieneś być powiedzieć mi prawdę” = „Nie mówiąc mi prawdy, złamałeś podstawowy (*general*) sąd powinnościowy – zasadę, pod którą się niniejszym podpisuję” = „□ powiedziałeś mi prawdę”.

§ 4.1.2 Analiza w języku teorii czynności mowy

Próba wyjaśnienia własności sądów moralnych poprzez odwoływanie się do imperatywów przyczyniła się jednak do interpretacyjnych nieporozumień. Między innymi zaczęto przypisywać Hare'owi re-

słówek „proszę” oraz „powinno się”. Zdanie, które w pracy tej zapisywane jest jako „! Zamykasz drzwi”, Hare zapisałby jako: „Zamykasz drzwi, proszę”.

¹³ Pojęcia uniwersalności pełnej i niepełnej zdefiniowane zostały w § 5.2.

¹⁴ Zob. § 5.1.

¹⁵ Por. tamże, s. 191-192.

dukcję zasad moralnych do imperatywów¹⁶, pomimo że według niego studiowanie nakazów jest jedynie najlepszym wstępem do analizy sądów moralnych¹⁷. Choć sądy moralne należą do tej samej rodziny wypowiedzi preskryptywnych co imperatywy¹⁸, nie znaczy to, iż są z nimi identyczne. Zalecenia posiadają również cechy, których nie mają imperatywy, co czyni je odrębną klasą wypowiedzi. Z tego względu Hare stopniowo odchodził od pierwotnego sposobu opisu elementu preskryptywnego. Zastąpił go analizą w języku teorii czynności mowy z charakteryzującymi ją technicznymi terminami, zaproponowanymi przez siebie samego, J. L. Austina oraz J. Searle'a.

Najbardziej rozpowszechnioną wśród filozofów teorię czynności mowy zaprezentował J. L. Austin: na początku w ramach teorii performatywów, a później teorii illokucji¹⁹. Najsłynniejszym, systematycznym jej rozwinięciem jest koncepcja czynności mowy J. Searle'a²⁰. Źródłem teorii illokucji była teoria performatywów Austina, której generalne wnioski przyjmował również Hare. Motywem jej powstania była chęć wykazania bezpodstawności przyjmowanego powszechnie założenia, nazwanego przez Austina **błędem opisowości**, który głosił, iż jedyną istotną funkcją zdań oznajmujących

¹⁶ Taką błędną interpretację uniwersalnego preskrytywizmu znaleźć można w: F. Ricken, *Etyka...*, dz. cyt., s. 33-34, oraz J. Harrison, *Dlaczego oceny moralne nie są imperatywami*, „Etyka” (22) 1986, s. 52. Harrison utożsamiał preskrypcje z imperatywami, twierdząc: „W myśl tego, co głosi profesor Hare, oceny moralne są nie tyle imperatywami, ile preskrypcjami, które pociągają za sobą imperatyw. Nie jestem przekonany co do tego, czy istotnie zachodzi jakakolwiek ważna poznawczo różnica między powiedzeniem, że są one imperatywami, a powiedzeniem, że są preskrypcjami”; tamże, s. 55. Zakładając dalej taką interpretację, Harrison odpowiada na pytanie zawarte w tytule powyższego artykułu. Odpowiedź jego można streścić w następujący sposób: uniwersalny preskrytywizm jest teorią błędną, ponieważ imperatywy nie mają wartości logicznych; nie można podstawić ich pod zmienne w rozumowaniach logicznych; nie mogą być błędne; nie można przedstawić jakichkolwiek racji na ich poparcie; niewypowiedziane imperatywy nie istnieją. Większość tych zarzutów jest wypadkową bardziej generalnego poglądu, iż imperatywy, w przeciwieństwie do sądów moralnych, nie mają charakteru przekonań. Jednak takie odczytanie uniwersalnego preskrytywizmu jest błędne, czego będziemy dowodzić poniżej.

¹⁷ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 2; G. H. Harman, R. M. Hare *and Moral Reasoning*, „Mind” (307) 1968, s. 427-428.

¹⁸ Zob. tamże, s. 3.

¹⁹ Zob. *Wypowiedzi performatywne oraz Jak działać słowami* w: J. L. Austin, *Mówienie...*, dz. cyt., s. 311-334, 545-708.

²⁰ Zob. J. Searle, *Czynności...*, dz. cyt.

jest funkcja opisowa²¹. Według niego, błędne traktowanie wielu wyrażań jako opisów jest spowodowane ich własnościami gramatycznymi. Deskrytywizm bowiem w różnych swoich odmianach rozpowszechnił osobliwe przekonanie, że wszystkie słowa są nazwami²². Gramatyka nie determinuje jednak własności logicznych. Austin podał w wątpliwość również przeświadczenie, iż wszystkie zdania sensowne są prawdziwe lub fałszywe, a wartości logiczne mogą mieć tylko zdania orzekające lub opisujące²³. Jak pisał,

założenie, że każde twierdzenie zmierza do tego, by być prawdziwym, jest nie tylko jałowe, ale ponadto można kwestionować, czy w ogóle tak jest. Zasada logiki: «Każde zdanie musi być prawdziwe lub fałszywe» zbyt długo funkcjonowała jako najprostsza, najbardziej przekonująca i najbardziej rozpowszechniona postać *błędu opisowości*. Pod jej wpływem filozofowie «na siłę» interpretowali wszystkie «zdania» na wzór twierdzenia, że jakaś rzecz jest czerwona, które wygłaszamy wtedy, gdy właśnie obserwujemy tę rzecz²⁴.

W związku z tym Austin podzielił wypowiedzi na **konstatacje** i **performatywy**. Według niego, o pierwszych z nich orzekamy prawdę lub fałsz i spełniają one takie funkcje jak stwierdzanie, wskazywanie, informowanie, potwierdzanie itp.²⁵ Performatywy²⁶ natomiast, choć są zdaniami sensownymi (nie są pytaniami, wykrzyknieniami czy rozkazami), nie mogą przybierać wartości logicznych i nic nie konstatają, mimo że możemy nazywać je „udanymi” lub „nieudanymi”, „fortunnymi” lub „niefortunnymi”²⁷. Ich funkcją nie jest spełnienie czynności, którą opisujemy jako „mówienie czegoś”. Typowym przykładem performatywu jest wypowiedź: „Niniejszym nazywam ten okręt Stefanem Batory”, będącą częścią procedury chrztu statku²⁸.

Z wprowadzonymi przez Austina dystynkcjami związane były jednak pewne problemy. Główną trudnością była niemożność wytyczenia jasnej granicy między performatywami a konstatacjami. Spostrzegłszy wielofunkcyjność czynności mowy, Austin zaproponował nową teorię, którą nazwać można „teorią illokucji”. Jest ona

²¹ Zob. J. L. Austin, *Mówienie...*, dz. cyt., s. 174.

²² Zob. tamże, s. 552.

²³ J. Szymura, *Język...*, dz. cyt., s. 115.

²⁴ J. L. Austin, *Mówienie...*, dz. cyt., s. 174.

²⁵ Zob. tamże, s. 678-691.

²⁶ Zob. np. tamże, s. 311-334, 553-556.

²⁷ Zob. tamże, s. 563. Wyszczególnienie różnych rodzajów niefortunności znaleźć można tamże, s. 561-598.

²⁸ Zob. tamże. 314.

bardziej ogólna od teorii performatywów i ją w sobie zawiera²⁹. Z jednej strony, koncepcja illokucji traktuje każdą poprawną wypowiedź jako **akt lokucyjny**, czyli posiadający sens i odniesienie, a więc zgodny albo niezgodny z faktami³⁰; z drugiej strony natomiast, bierze pod uwagę to, że każda taka wypowiedź jest również **aktem illokucyjnym**, co oznacza, iż może być traktowana jako udana albo nieudana z uwagi na dotyczące ją konwencje³¹. Dawne odróżnienie performatywy–konstatacje było odróżnieniem dzielącym zbiór wypowiedzi na dwa rodzaje – jedne były robieniem czegoś, drugie – tylko mówieniem. Teoria illokucji nie wprowadza takiego podziału, lecz w ramach każdej poprawnej wypowiedzi wyodrębnia trzy jej aspekty. **W aspekcie lokucyjnym** wypowiedź jest mówieniem czegoś. **W aspekcie illokucyjnym** jest ona tym, co się robi mówiąc – konwencjonalnym³² działaniem dokonywanym za pomocą aktu lokucyjnego, np. obiecywaniem, nakazywaniem, pytaniem, proszeniem itd.³³ Trzecim aktem zasługującym na wyróżnienie jest **akt perlokucyjny**, który polega na tym, że poprzez powiedzenie czegoś (i wykonanie aktu illokucji) czynimy coś jeszcze innego (akt perlokucji), np. nakłonienie kogoś do pewnych postaw i uczuć³⁴. Tak więc lokucją wypowiedzi: „Czy możesz zamknąć drzwi?” jest jej sens i odniesienie zawarte również w zdaniu „Możesz zamknąć drzwi”. Mocą illokucyjną pierwszego z nich jest pytanie, drugiego zaś stwierdzenie, natomiast w aspekcie perlokucyjnym mogą one powodować dość nieprzewidywalne skutki, np. zamknięcie drzwi, trzaśnięcie drzwiami, otworzenie ich jeszcze szerzej itd.

§ 4.2 Czynności mowy a znaczenie

Większość teorii czynności mowy³⁵ akceptuje przedstawiony powyżej klasyczny podział na trzy rodzaje czynności językowych. Rów-

²⁹ Por. M. Hempoliński, *Brytyjska...*, dz. cyt., s. 123. Zob. J. L. Austin, *Jak działać słowami* w: tenże, *Mówienie...*, dz. cyt., s. 545-708.

³⁰ Zob. J. L. Austin, *Mówienie...*, dz. cyt., s. 640-641.

³¹ Zob. tamże, s. 644-646.

³² Zob. tamże, s. 652-653.

³³ Por. J. Szymura, *Język...*, dz. cyt., s. 213.

³⁴ Zob. J. L. Austin, *Mówienie...*, dz. cyt., s. 646-648.

³⁵ Koncepcje aktów mowy oprócz Austina rozwijali m.in. K. Bach, H. P. Grice, R. M. Hare, R. M. Harnish, J. Searle, P. F. Strawson, S. L. Tsohatzidis.

niez R. M. Hare posługiwał się takimi pojęciami, jak lokucja, illokucja czy perlokucja. Między poszczególnymi przedstawicielami teorii czynności mowy, takimi jak J. L. Austina, P. Strawsona, J. Searle'a i H. P. Grice'a, istniały jednak również rozbieżności. Spowodowane były one różnym pojmowaniem m.in. relacji pomiędzy znaczeniem wypowiedzi a jej mocą illokucyjną (znaczeniem wyrażenia i mocą illokucyjną zdania, które je zawiera) oraz odmiennym rozumieniem roli, jaką dla znaczenia wypowiedzi odgrywają intencje mówiących i konwencje językowe. Przedstawiono również zarzuty przeciwko teorii czynności mowy. Są one dla nas ważne, gdyż większość argumentów krytycznych, które kierowano w stronę tej teorii, dotyczy także metaetyki Hare'a.

Richard Hare zaliczał siebie do filozofów, których nazwał performerami (*performers*)³⁶. **Performerzy** uważają, że znaczenie niektórych słów może być wyjaśnione, lub przynajmniej częściowo wyjaśnione, poprzez wskazanie, iż zawarcie danego słowa w odpowiednim zdaniu na odpowiednim miejscu nada całemu temu zdaniu taką własność, iż wypowiedzenie go w odpowiednim kontekście będzie dokonaniem konkretnej czynności illokucyjnej³⁷. Hare uważał, że powyższej metody analizy znaczenia bronić można nie tylko w odniesieniu do słów i zdań moralnych, lecz w stosunku do całego języka³⁸. John Searle, krytykując performerów, ich sposób myślenia zrekonstruował w następujący sposób³⁹:

Aby dowiedzieć się, jakie znaczenie ma słowo D (np. „dobry”), performerzy stawiają pytanie:

(1) Jak używa się D?

Następnie przyjmują milcząco, że pytanie (1) znaczy:

(2) Jak używa się D w prostych, kategoriowych zdaniach oznajmujących w czasie teraźniejszym o postaci „X jest D”?

Aby rozstrzygnąć (2), zakładają następnie, że rozumieć je należy jako pytanie:

³⁶ Zob. R. M. Hare, *Meaning and Speech Acts*, w: R. M. Hare, *Practical Inferences...*, dz. cyt., s. 75.

³⁷ Zob. tamże, s. 75.

³⁸ Krytykę definiowania znaczenia poprzez akty mowy znaleźć można w: G. J. Warnock, *Hare on Meaning and Speech Acts*, „The Philosophical Review” (1) 1971. Odpowiedź Hare'a znajduje się w: R. M. Hare, *Appendix. Reply to Mr. G. J. Warnock*, w: tenże, *Practical Inferences...*, dz. cyt., s. 95.

³⁹ Zob. J. Searle, *Czynności...*, dz. cyt., s. 186-187.

(3) Jak używa się zdań zawierających D?

A to z kolei pojmują jako:

(4) Jaką czynność illokucyjną wykonuje się wypowiadając takie zdania?

Bądź

(5) Jakie są warunki wygłaszania niewadliwych twierdzeń, gdy wypowiada się takie zdania? (Kiedy faktycznie powiedzielibyśmy coś o postaci: „X jest D”?).

Powyższą analizę rozumieć można na dwa sposoby⁴⁰. Pierwszy polega na założeniu, że kiedy performerzy mówią „D używa się, by wykonać czynność C”, uważają, iż każde literalne wypowiedzenie słowa D jest spełnieniem czynności C. Mniej radykalni przedstawiciele tego podejścia, w tym również Hare⁴¹, często łagodzą swoją konkluzję, twierdząc, że „podstawowym użyciem D jest wykonanie czynności C”.

§ 4.3 Zarzut błędu redukcji sądów moralnych do zaleceń

John Searle uważał, że nie można wyjaśnić znaczenia słowa „dobry” przez powiedzenie, iż używa się go do wykonywania czynności chwaleń⁴². Można podać nieskończenie wiele zdań, w których słowo „dobry” występuje w sposób literalny, a jednak wypowiadając owe zdania, nie spełniamy takiej czynności mowy. Np. zdanie „Jeśli to jest dobre, powinniśmy to kupić” nie jest równoważne zdaniu: „Jeśli to chwałę, powinniśmy to kupić”; „To jest zazwyczaj dobre”, nie jest równoważne zdaniu „Zazwyczaj to chwałę”; „Zastanawiam się, czy to jest dobre” nie jest równoważne „Zastanawiam się, czy to chwałę” itd.⁴³ W podobny sposób, zakładając tym razem, że Hare utożsamiał oceny z zaleceniami, w artykule *Evaluative and Descriptive* atakował go również G. M. Matthews⁴⁴. Starał się on pokazać, iż wszystkie próby sprowadzenia sensu ocen do zaleceń powodują nieakceptowalne konsekwencje. Powyższe argumenty opierają się jednak na błędnym rozumieniu teorii znaczenia, którą posługiwał się

⁴⁰ Zob. tamże, s. 175.

⁴¹ Zob. R. M. Hare, *Meaning...*, dz. cyt., s. 77.

⁴² Por. J. Searle, *Czynności...*, dz. cyt., s. 174.

⁴³ Por. tamże, s. 177.

⁴⁴ Zob. G. M. Matthews, *Evaluative and Descriptive*, „Mind” (67) 1958, s. 335.

Hare. Krytyka Searle’a i Matthews’a jest bowiem zasadna tylko wtedy, gdy sądy zredukują się do zaleceń, nakazów czy imperatywów oraz przypisze się Hare’owi akceptację tezy, że każde literalne wypowiedzenie słowa moralnego jest spełnieniem czynności mowy o charakterze preskrypcji. Jednak Hare nie akceptował założenia, że „x jest dobre” = „zalecam/nakazuję x”. Według niego, wyrażenia te nie są wymienialne *salva veritate*⁴⁵. Gdybyśmy powiedzieli, że znaczenie słowa „dobry” sprowadza się do mocy illokucyjnej, byłoby to nadmierne uproszczenie. Zdaniem Hare’a, performer nie ma tego na myśli, kiedy mówi, że jest to słowo używane do zalecania; twierdzi jedynie, że w zdaniach ze słowem „dobry” zamiast mocy illokucyjnej deskrypcji standardowo występuje moc illokucyjna preskrypcji⁴⁶. Co więcej, Hare nigdy nie twierdził, że znaczenie terminów moralnych możemy wyjaśnić jedynie poprzez ich preskryptywność. To, że oceny mają cechy podobne do nakazów czy pochwał, jest warunkiem koniecznym ich tożsamości jako ocen, nie jest to jednak warunek wystarczający. Znaczenie sądów moralnych jest bowiem wyznaczone przez koniunkcję wszystkich formalnych praw, które tworzą logikę moralności. W świetle jego teorii również uniwersalizowalność⁴⁷, posiadanie uzasadnienia⁴⁸ czy nadrzędność⁴⁹ wnosi do znaczenia niezbędne cechy. To, że znaczenie sądów moralnych jest zdeterminowane wieloma własnościami formalnymi, których złożenie jest specyficzne jedynie dla nich, powoduje, iż nie można podać w języku potocznym prostego ich odpowiednika.

§ 4.4. Zarzut błędu czynności mowy

Zdaniem Searle’a analiza znaczenia w terminach czynności mowy, nawet jeśli nie miałaby być prostą redukcją sądów moralnych do zaleceń, ostatecznie jednak i tak jest błędna, gdyż znaczenie słowa i moc illokucyjna zawierającego je zdania to dwie odrębne rzeczy⁵⁰. Znaczenie wyznaczone jest nie przez akt illokucyjny, lecz lokucyjny

⁴⁵ Zob. R. M. Hare, *Meaning...*, dz. cyt., s. 77.

⁴⁶ Zob. tamże, s. 93.

⁴⁷ Zob. § 5.2.

⁴⁸ Zob. § 5.1.

⁴⁹ Zob. § 4.14.

⁵⁰ Por. J. Searle, *Czynności...*, dz. cyt., s. 175.

(sens i odniesienie). Pomimo że „czynności mowy, czyli czynności wykonywane, gdy wypowiada się zdanie, są w ogólności funkcją znaczenia tego zdania”⁵¹, to jednak funkcja ta nie jest tożsama ze znaczeniem danej wypowiedzi⁵². „Znać znaczenie terminu ogólnego, a więc wyrażenia orzecznikowego, to wiedzieć, w jakich warunkach jest on prawdziwy czy fałszywy o danym przedmiocie”⁵³. Twierdzenie, że odpowiedzi na pytanie (4) (Jak używa się zdań zawierających D?) przynoszą nieuchronnie rozstrzygnięcie znaczenia słowa D, nazwać należy „**błędem czynności mowy**”⁵⁴. Co więcej, taka analiza nie dostarcza nam „żadnego sposobu odróżniania użycia słowa od użycia zawierającego je zdania”⁵⁵.

Odpowiedź na zarzut błędu czynności mowy była dla Hare’a bardzo ważna, oparł on bowiem swój nondeskrytywizm na założeniu, iż znaczenie sądów i terminów moralnych wyznaczone jest m.in. przez moc illokucyjną zalecenia. W esejach *Meaning and Speech Acts*⁵⁶ oraz *Austin’s Distinction between Locutionary and Illocutionary Acts*⁵⁷ próbował uzasadnić tezę, iż moc illokucyjna może być elementem znaczenia. Według niego, wyrażenia można podzielić na co najmniej trzy grupy ze względu na udział mocy illokucyjnej w ich znaczeniu. Można wyróżnić słowa, których:

- 1) znaczenie jest niezależne od mocy illokucyjnej (nazwijmy je **czysto deskryptywnym**);
- 2) znaczenie wyznaczone jest zarówno przez lokucję jak i illokucję;
- 3) znaczenie wyznaczone jest jedynie przez illokucję (do tej grupy należą słowa o znaczeniu **czysto ewaluatywnym**)⁵⁸.

Do grupy pierwszej zaliczają się takie pojęcia, jak koń, zielony, wybuch; do drugiej „gęste pojęcia moralne” (np. mężny, hojny, odważny, uczciwy)⁵⁹; do trzeciej zaś „chude pojęcia moralne” (np. dobry, zły, powinno się) oraz czasowniki wyraźnie illokucyjne (np.

⁵¹ Tamże, s. 31.

⁵² Zob. tamże, s. 162.

⁵³ Tamże, s. 162. Por. Hare’a definicję deskrytywizmu (§ 3.2).

⁵⁴ Por. tamże, s. 186-187.

⁵⁵ Tamże, s. 186.

⁵⁶ R. M. Hare, *Meaning...*, dz. cyt., s. 74-93.

⁵⁷ Tenże, *Austin’s Distinction between Locutionary and Illocutionary Acts*, w: *tenże, Practical Inferences...*, dz. cyt., s. 100-114.

⁵⁸ Zob. tamże, s. 110.

⁵⁹ Zob. tenże, *Freedom...*, dz. cyt., s. 27.

obietuję, stwierdzam, pytam). Terminy etyczne należące do grupy (2) składają się ze znaczenia ewaluatywnego i deskryptywnego; będące zaś w grupie (3) – jedynie z ewaluatywnego⁶⁰.

O istocie znaczenia nakazowego, **znaczenia ewaluatywnego**⁶¹ (ZE) nie stanowi funkcja emotywna czy ewokatywna (jak w wypadku znaczenia emotywnego u Stevensona⁶²), lecz przede wszystkim funkcja preskryptywna – konwencjonalnie wyznaczona moc illokucyjna zalecenia. To właśnie znaczenie ewaluatywne jest owym „czymś dodatkowym”, co zdaniem Hare’a wraz z warunkami prawdziwości i składnią współtworzy znaczenie sądów moralnych⁶³. Jest ono **znaczeniem pierwotnym** (*primary meaning*), ZE determinuje bowiem cechy istotne dla pojęć i sądów moralnych, nawet jeśli w pewnych kontekstach zostały one użyte jedynie w funkcji opisowej⁶⁴. Znaczenie sądów moralnych jest ewaluatywnie pierwotne, lecz nie każde ich użycie musi być ewaluatywne.

Znaczenie deskryptywne (ZD) natomiast odpowiada za element semantyczny wypowiedzi moralnej⁶⁵. Jest ono **znaczeniem wtórnym** (*secondary meaning*), gdyż nie stanowi o istocie sądów moralnych, choć, co odróżnia je od imperatywów, może być użyte jako ich podstawowe znaczenie⁶⁶. ZD jest ściśle związane z uniwersalizowalnością sądów moralnych. Deskryptywność Hare przedstawiał dwojako⁶⁷: po pierwsze, jako cechę, która powoduje, że np. sąd „Ten człowiek jest dobry” dostarcza czysto opisowych informacji o danym człowieku (np. nie krzywdzi on innych, lecz im pomaga itd.); po drugie, jako informację o warunkach prawdziwości danego sądu. Owe aspekty deskryptywności są ze sobą ściśle powiązane. Ostatecznie należałoby się zgodzić z G. M. Matthewsem⁶⁸, że pojęcie znaczenia

⁶⁰ Trudno jest jednoznacznie stwierdzić, czy Hare zaliczyłby wszystkie czasowniki wyraźnie illokucyjne do grupy (2), czy (3). Musimy bowiem pamiętać, że znaczenie zdań, które je zawierają, wyznaczone jest również przez aspekt lokucyjny. Uczynił tak na pewno z chudymi pojęciami moralnymi (dobry, powinno się). Zob. tamże, s. 22; tenże, *Myślenie...*, dz. cyt., s. 14.

⁶¹ Zob. tamże, s. 35.

⁶² Zob. Ch. L. Stevenson, *Ethics...*, dz. cyt., s. 37-62. Zob. § 2.2.

⁶³ Zob. § 3.2.

⁶⁴ Zob. R. M. Hare, *Freedom...*, dz. cyt., s. 22.

⁶⁵ Zob. tenże, *Sorting...*, dz. cyt., s. 6.

⁶⁶ Zob. tenże, *Imperatives, Prescriptions...*, dz. cyt., s. 51.

⁶⁷ Por. tenże, *The Language...*, dz. cyt., s. 109-126.

⁶⁸ Zob. M. Matthews, *Evaluative...*, dz. cyt., s. 337.

deskryptywnego Hare stosował bardzo liberalnie. Słusznie zauważył, że Hare używał go do określenia każdej informacji, która może być przenoszona np. przez słowo „dobry”. Najczęściej tą informacją są standardy użycia danego wyrażenia w określonym kontekście oraz racje przypisania predykatu moralnego ocenianej rzeczy – zarówno subiektywne, jak i obiektywne, gdyż związane z namysłem idealnego nakazodawcy⁶⁹. Istnieje zależność pomiędzy informacją a kryteriami poprawnego użycia słów, jak również informacją a świadectwem na rzecz danej oceny⁷⁰. Znać deskryptywne znaczenie wypowiedzi moralnej to bowiem wiedzieć, za pomocą jakich standardów nadawca tej wypowiedzi osądza dane czyny⁷¹. Informacyjność, a zatem również deskryptywność naszych wypowiedzi wzrasta wraz ze wzrostem powszechności i oczywistości wiedzy dotyczącej tego, jaki zbiór cech naturalnych stał się podstawą oceny danego przedmiotu czy czynu. Jeśli jakaś ocena danego przedmiotu A jest społecznie „szywno” skorelowana z jakimś zbiorem istotnych, relewantnych własności naturalnych *RNatA* (świadectwem), to należy powiedzieć, iż wypowiedź taka jest pierwotnie deskryptywna, a w niektórych przypadkach wręcz czysto deskryptywna. Jeżeli w zdaniu „To jajko jest dobre” ocena jajka A jako *Mor1* (dobre) dokonuje się tylko i wyłącznie wtedy, gdy każdy kompetentny użytkownik języka przypisuje mu cechę „niebycia zepsutym” (*RNatA*), to powiedzieć można, iż zdanie: „To jajko jest dobre” znaczy: „To jajko nie jest zepsute” (*Mor1A* = *RNatA*). Wypowiedź taka byłaby czysto deskryptywna.

Zdaniem Hare’a, ewaluatywność jest parametrem dyskretnym – binarnym, deskryptywność zaś parametrem ciągłym. Dana wypowiedź nie może być bardziej lub mniej normatywna: albo użyta jest ona ewaluatywnie, albo nie. Odwrotnie jest jednak z deskryptywnością⁷². Pewne wyrażenia są bardziej deskryptywne, inne natomiast mniej. Te dziedziny, w których oceny faktów czy przedmiotów są oparte na powszechnie podzielanych racjach, cechuje większy stopień deskryptywności, gdyż niosą one ze sobą więcej pewnych in-

⁶⁹ Zob. § 6.3.3.

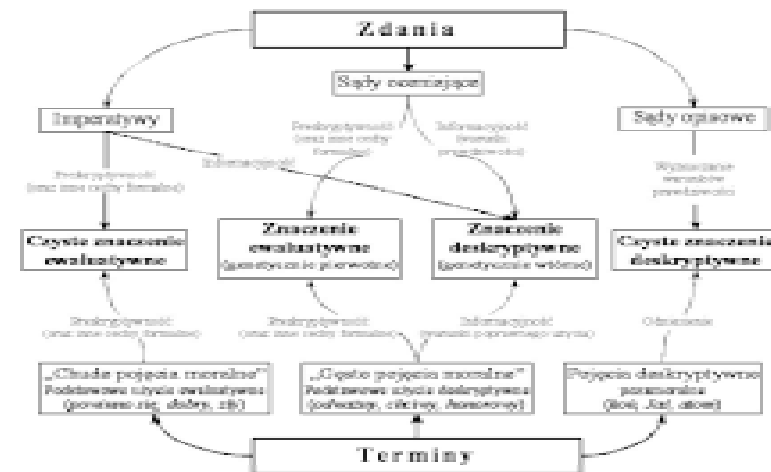
⁷⁰ Por. R. M. Hare, *The Language...*, dz. cyt., s. 116.

⁷¹ Zob. tamże, s. 146. Krytykę Hare’a ujęcia znaczenia deskryptywnego znaleźć można w: G.N.A. Versey, *Hare on Descriptive Meaning Rules*, „Mind” (296) 1965, s. 590-591.

⁷² Por. R. M. Hare, *The Language...*, dz. cyt., s. 122.

formacji. Jeśli powszechna zgoda co do kryteriów oceny jest bardzo duża, istnieje niekiedy pokusa lub zachodzi fakt redukcji normy do opisu. W innych dziedzinach, w których kryteria zmieniają się z biegiem czasu bądź praktyki oceniające mają krótką historię, oceny pozostają **pierwotnie ewaluatywne**. Przykładem pierwszej są nauki ścisłe (dobry dowód matematyczny, dobra teoria, dobry eksperyment, dobra obserwacja), drugiej zaś estetyka (dobry wiersz, dobra muzyka itd.). Język moralny posiada cechy dwóch powyższych dziedzin. Zbiór sądów moralnych rozpada się na podklasę sądów o stabilnym znaczeniu deskryptywnym i tych, które są powodem wielu w tym względzie kłopotów. Sąd „Zabijanie bezbronnych, niewinnych dzieci dla zabawy jest złe” ma bowiem inny charakter niż sąd „Klonowanie ludzi jest złe”. Główna różnica pomiędzy nimi polega na tym, iż warunki prawdziwości pierwszego z nich są społecznie uzgodnione i raczej powszechnie akceptowane, drugi natomiast ciągle wymaga świeżego namysłu na poziomie krytycznym.

Relacje pomiędzy różnymi rodzajami znaczenia graficznie przedstawić można w następujący sposób:



Rysunek 4. Rodzaje znaczeń

Uznawana przez Hare'a dychotomia: znaczenie ewaluatywne–deskrytywne nie jest jednak powszechnie akceptowana⁷³. Podstawowy zarzut uderza w niejasność kryteriów, które pozwalałyby rozstrzygać, czy dany termin przynależy do grupy (1), (2), czy (3). Zdaniem Philipa Foot⁷⁴, trudno znaleźć bezsporne warunki, które jednoznacznie odpowiedzą nam, dlaczego wyraz „dobry” czy „powinno się” ma podstawowe znaczenie ewaluatywne, a „hojny” lub „samolubny” – podstawowe znaczenie deskrytywne, bez zakładania koncepcji znaczenia Hare'a. Głównym problemem, z którym poradzić musi sobie preskrytywizm, jest uzasadnienie odpowiedzi na pytanie, czy w ogóle jest możliwe wydobycie owego elementu nazwanego „znaczeniem ewaluatywnym” ze znaczenia takich słów jak „dobry”. Ponieważ zwroty oceniające zawsze występują razem z wyrazami, do których się odnoszą, nie można sensownie mówić o znaczeniach takich słów w oderwaniu od ich przedmiotu. „Chude pojęcia moralne” nabierają znaczenia zawsze w odniesieniu do obiektów, które za ich pomocą są oceniane. Możemy więc wiedzieć, co znaczy „dobry człowiek” i „dobry koń”, jednak nie możemy pytać, co znaczy „dobry jako taki”⁷⁵.

§ 4.5 Odpowiedź Hare'a na argument błędu czynności mowy

Rozwiązanie powyższych problemów przynieść może zanalizowanie ich w języku czynności mowy, zdaniem Hare'a bowiem nieostrość dychotomii znaczenie deskrytywne–ewaluatywne jest spowodowana nieostrością podziału na znaczenie i moc illokucyjną. Co więcej,

⁷³ Zbiór argumentów krytycznych skierowanych w stronę dychotomii znaczenie ewaluatywne–deskrytywne Czytelnik może znaleźć w: M. Forrester, *An Argument For Descriptivism*, „The Journal of Philosophy” (20) 1974, s. 759-769. Spotkać można również filozofów, którzy bronili tego rozróżnienia. W tej kwestii tropem Stevenson'a i Hare'a starał się podążać J. O. Urmson; zob. J. O. Urmson, *The Emotive Theory of Ethics*, Oxford University Press, New York 1969, s. 138-146. Jednak także jego próby uczynienia dychotomii ewaluatywne – deskrytywne bardziej wyraźną poprzez zaprezentowanie jej w języku teorii aktów mowy spotykały się z krytyką; zob. L. J. Jost, *On a Prescriptivist Dichotomy*, „Mind” (338) 1976, s. 258-304.

⁷⁴ Zob. P. Foot, *Goodness and Choice*, w: W. D. Hudson (ed.), *The Is-Ought Question*, Macmillan, London 1969, s. 214-239.

⁷⁵ Zob. tenże, *Moral Beliefs*, w: W. D. Hudson (ed.), *The Is-Ought...*, dz. cyt., s. 196. Odpowiedź na ten zarzut znaleźć można w § 4.10.

taka nieostrość nie falsyfikuje tezy, iż istotą sądów moralnych jest znaczenie ewaluatywne – w skład ich znaczenia wchodzi moc illokucyjna preskrypcji. To właśnie sztucznie wyostrzona dystynkcja znaczenie – moc illokucyjna stanowi o błędzie deskrytywistów⁷⁶.

Hare zadał retoryczne pytanie:

„Czy o kimś, kto nie wie, że człowiek wypowiadający «Obiecuję», obiecuje, lub człowiek mówiący «Kot jest na macie», wyraża stwierdzenie, powiedzieliśmy, że zna znaczenie tych słów? Odpowiedź wydaje się z pewnością negatywna. [...] Jakikolwiek pełne wyjaśnienie znaczenia wyrazu pojawiającego się w zdaniu musi zawierać wyjaśnienie jego trybu gramatycznego, jak również jego czasu, osoby, strony itd.; trudno wyobrazić sobie, jak mogłoby to być zrobione inaczej niż poprzez sprecyzowanie rodzaju aktu mowy, do którego dany tryb gramatyczny jest przypisany przez konwencje konstytuujące język⁷⁷.”

Według Hare'a, sens słowa „jest” w zdaniu „Kot jest na macie” wyznacza zarówno pojęcie predykcji, jak i pojęcie stwierdzenia (*assertion*). Mówi nam ono bowiem, że zdanie to nie jest rozkazem ani pytaniem. Z tego też względu pojęcie sensu jako składnik lokucji, przynajmniej niekiedy (jeśli nie zawsze), zawiera w sobie konkretną moc illokucyjną⁷⁸. Potwierdzenie oraz obronę tego stanowiska można znaleźć w rozprawie J. Szymury⁷⁹. Według niego, w niektórych przypadkach znaczenie jest zależne od mocy illokucyjnej. Choć nie odniósł się on bezpośrednio do Hare'a, analizując Austina teorię czynności mowy, stwierdził, że

w myśl bronionego na tych stronach poglądu zdania wyraźnie illokucyjne, w których skład wchodzi illokucyjne czasowniki, mają ten charakter, że co najmniej częścią tego, co się w nich stwierdza, jest stwierdzenie dokonania określonej illokucji przez osobę takie zdanie wygłaszającą. Tym samym trzeba uznać, iż przynajmniej w przypadku tego typu wyrażań moc illokucyjna jest zawarta w znaczeniu. Kiedy ktoś mówi «Obiecuję, że p» między innymi znaczy to, że obiecuje i taka właśnie jest illokucyjna funkcja tych słów. Austińskie odróżnienie aktu illokucyjnego od aktu lokucyjnego jako zasadniczo odmiennych, jeżeli nawet da się utrzymać, nie jest tak ostre, jak się to autorowi *How to Do Things with Words* wydawało. [...] Ktoś, kto mówi *Obiecuję, że będę w domu o siódmej*, obiecuje dlatego, że użyte przez niego zdanie znaczy to, co właśnie znaczy. Innymi słowy, osoba, o której mowa, robi coś, o czym mówi, że to robi i co jest znaczeniem tego, co mówi. [...] Tak więc illokucyjna

⁷⁶ Zob. R. M. Hare, *Austin's Distinction...*, dz. cyt., s. 113.

⁷⁷ Zob. tenże, *Meaning...*, dz. cyt., s. 78.

⁷⁸ Zob. tenże, *Austin's Distinction...*, dz. cyt., s. 107.

⁷⁹ Szczegółowe przedstawienie relacji pomiędzy znaczeniem a mocą illokucyjną wnikliwie przedstawia J. Szymura; zob. J. Szymura, *Język...*, dz. cyt.

moc wyraźnie illokucyjnego wyrażenia jest w pewnym sensie czymś, co to wyrażenie między innymi znaczy. Dlatego jest elementem jego znaczenia. To, co można powiedzieć o niej jako o elemencie znaczenia, nie jest jednak tym samym, co można o niej powiedzieć właśnie jako o *illocutionary force*⁸⁰.

§ 4.6 Znaczenie a kryterium użycia

Dychotomia znaczenie–kryterium (znaczenie słowa/wypowiedzi–kryteria jego użycia/racje stojące za daną oceną) jest kluczową nondeskryptywistyczną tezą Hare’a. Uniknął on dzięki niej tego, co za Austinem nazwał błędem opisowości⁸¹. Jego zdaniem popełniają go deskryptywiści (np. Phillipa Foot), uważają bowiem, że znać znaczenie słowa tworzącego dyskurs moralny to tyle, co znać kryteria jego użycia – wiedzieć w odniesieniu do czego możemy je zastosować. Jednak według Hare’a, znaczenie słów i wypowiedzi moralnych jest pierwotnie ewaluatywne i to stanowi o ich istocie. Rozdzielenie znaczenia od kryteriów użycia Hare uzasadniał, opierając się na analizie języka potocznego, w szczególności na takich fenomenach lingwistycznych, jak możliwość stosowania słów pierwotnie ewaluatywnych do nowych klas przedmiotów, istnienie i charakter reform moralnych, realność konfliktów moralnych i możliwość wzajemnego zrozumienia oponentów podczas sporu moralnego.

Zdaniem Hare’a, już G. E. Moore stwierdził, że czym innym jest znaczenie słowa „dobry”, a czym innym opis zbioru przedmiotów, które tę własność posiadają⁸². Sprowadzanie predykatów moralnych do opisu kryteriów ich użycia oznaczałoby więc

traktowanie ich tylko jako konwencji językowych, obowiązujących [...] tych, którzy chcą poprawnie mówić danym językiem. Błąd polegający na myleniu konwencji moralnych z konwencjami językowymi jest analogiczny do błędu, który polega na myleniu intuicji moralnych z językowymi⁸³.

Konwencje moralne odpowiadają za korelowanie konkretnych zbiorów istotnych własności naturalnych *RNat* z danymi ocenami *Mor*, konwencje

⁸⁰ Tamże, s. 235, 254, 255.

⁸¹ **Błąd opisowości** (*descriptive fallacy*) – nieuwzględnianie nieopisowych funkcji słów, traktowanie ich, gdy jest to niezgodne z powszechnymi zwyczajami językowymi, jako opisów. Konsekwencją takiego błędu jest redukcja różnych sposobów użycia słów do deskrypcji. Por. tenże, *Czy rozmowy o prawdzie są bezprzedmiotowe?*, „Kwartalnik Filozoficzny” (2) 1995, s. 24.

⁸² G. E. Moore, *Zasady...*, dz. cyt., s. 13-14.

⁸³ R. M. Hare, *Myślenie...*, dz. cyt., s. 92.

lingwistyczne natomiast za to, iż takie pojęcia, jak „dobry” czy „powinno się”, służą do wyrażania uniwersalnych, nadrzędnych zaleceń.

Według Hare’a, znaczenie ewaluatywne jest pierwotne, gdyż tylko ono pozostaje niezmiennie względem kontekstów użycia, znaczenie deskryptywne natomiast zależy od kryteriów, jakie w danym momencie oraz w danej społeczności są uznane za obowiązujące. Jedynie czysta „własność” ewaluatywna *Mor*, wyrażona za pomocą „chudych pojęć moralnych”, będących *de facto* operatorami logiki moralnego użycia słów, znaczeniowo pozostaje niezmienna i może zostać przypisana różnym zbiorom własności *RNat*. Owe „chude pojęcia moralne”, jako operatory niezmiennej logiki moralności, mogą być związane z różnymi przedmiotami, ze względu na posiadane przez nie własności naturalne. Znaczenia ewaluatywnego nie można zredukować lub choćby jednoznacznie skorelować z jakimś jednym zbiorem własności naturalnych, gdyż nie istnieje taki zbiór, który byłby powszechnie akceptowaną racją za przyjęciem danej oceny *Mor*. Zasadnicza różnica, która powoduje niezgodę pomiędzy naturalistką Phillipą Foot a Hare’em polega na tym, iż według Hare’a nie możemy, w co wierzy Foot⁸⁴ czy J. Griffin⁸⁵, znaleźć jednego zbioru własności *RNat*, który mógłby pełnić rolę warunków prawdziwości danej oceny. Hare, podobnie jak Kant, przeciwstawiał się tezie, że etykę można oprzeć na psychologicznych sądach o naszej naturze. Natura człowieka jako jeden i absolutny zbiór przekonań i pragnień nie istnieje, ludzie nie mają jednego celu, do którego jednomyślnie dążą. Nawet jeśli zgodzimy się, że może być nim szczęście czy wzrastanie, to dla każdego z nas co innego one znaczą. Nie możemy wiedzieć nic na temat tego, czego „człowiek jako taki” uniwersalnie chce (*need*), jakie ma potrzeby i co zaspokoi jego preferencje. Hare przeczył, że człowiek jako powszechnik i przypisywane mu na drodze apriorycznej cechy wyczerpują to, czego indywidualni, realni ludzie mogą pragnąć. Przypisanie ludziom takich cech na drodze aposteriorycznej byłoby natomiast jedynie statystyczną generalizacją, której nie możemy używać jako bazy myślenia moralnego⁸⁶. Uprawiając ety-

⁸⁴ Zob. P. Foot, *Goodness...*, dz. cyt., s. 216.

⁸⁵ Zob. J. Griffin, *Interpersonal Comparability*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., 84-88.

⁸⁶ Zob. R. M. Hare, *Comments on Griffin*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 240-241.

kę, możemy polegać jedynie na naszej umiejętności wybierania celów i praktycznej zdolności godzenia ze sobą skonfliktowanych preferencji. Sednem moralności musi być racjonalna, zgodna z językiem moralnym, koordynacja ludzkich interesów. Ów stan rzeczy jest jedną z przyczyn niemożności zredukowania znaczenia ewaluatywnego do deskryptywnego.

W dwóch różnych społecznościach w zdaniu „To jest dobry człowiek” znaczenie ewaluatywne zawsze pozostanie niezmiennie, znaczenie deskryptywne zaś może być różne w zależności od przyjmowanych kryteriów bycia dobrym człowiekiem⁸⁷. Gdybyśmy nie przeprowadzili dychotomii znaczenie–kryterium, nawet jeśli znalibyśmy znaczenie wyrażenia „dobry kij do krykieta” czy „dobry chronometr”, nie potrafilibyśmy użyć słowa „dobry” z nowymi klasami obiektów⁸⁸. Możemy jednak mówić znacząco o „dobrej gaśnicy” czy „dobrym samochodzie”, a nawet „dobrym Marsjaninie”. Zatem skoro rozumiemy znaczenie i sensownie używamy słowa „dobry” we wszystkich kontekstach (również wtedy, gdy kryteria prawdziwości sądów je zawierających są różne, a nawet nieznanne), znaczenie owego słowa we wszystkich przypadkach musi być takie samo.

Kolejnym powodem uprzywilejowanego traktowania znaczenia ewaluatywnego jest fakt, że używane jest ono do tego, aby zmienić znaczenie deskryptywne⁸⁹. Dzieje się tak często podczas reformy obowiązującego etosu. Z tego powodu odrzucenie przez chrześcijan moralności greckiej i żydowskiej nie przerwało racjonalnego dialogu między uczestnikami owych kultur, nie spowodowało niewspółmierności pojęciowej na różnych etapach historii. W procesie dziejowym zmianie ulegają bowiem kryteria stosowania słów moralnych do oceny danych czynów, lecz nie ich pierwotne znaczenie ewaluatywne. Tacy reformatorzy jak Jezus za pomocą słów „dobry” lub „powinno się”, których znaczenie było powszechnie znane, mogli zalecać zmianę kryteriów ich stosowania⁹⁰. Powyższą hipotezę możemy uznać za uzasadnioną, gdy zdamy sobie sprawę z sensowności poniższego zdania oraz tego, co czyni ktoś, mówiąc:

⁸⁷ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 118.

⁸⁸ Por. tamże, s. 96, oraz tenże, *Myslenie...*, dz. cyt., s. 97.

⁸⁹ Zob. tenże, *The Language...*, dz. cyt., s. 119.

⁹⁰ Zob. tenże, *Myslenie...*, dz. cyt., s. 92.

Choć każdy z was uważa, że dobrym człowiekiem jest ten, kto pomści swą krzywdę doznaną od wroga i ukamieniuje cudzołożnicę, powiadam wam, naprawdę dobrym człowiekiem jest ten, kto im wybaczy.

Zdaniem Hare’a, w obu przypadkach użycie określenia „dobry”, choć powodowane różnymi kryteriami poprawnego zastosowania, posiada takie samo znaczenie. Pomimo iż argument ten nie jest konkluzywny, gdyż podać można drugą interpretację powyższej wypowiedzi:

Choć każdy z was uważa, że wyrażenie „dobry człowiek” znaczy „B”, powiadam wam, że wyrażenie „dobry człowiek” naprawdę znaczy „C”,

wydaje się, iż wyjaśnienie Hare’a tego typu zdań jest bardziej intuicyjne, prostsze, zatem trafniejsze⁹¹.

W ostatnich latach teorię znaczenia, która oddziela znaczenie od kryterium, rozwijają T. Horgan i M. Timmons⁹², nazywając ją **wariantystyczną koncepcją znaczenia** (*variantist concept of meaning*). Według nich, wariantyzm⁹³ jest przekonaniem, że jedno i to samo pojęcie może być użyte przez dwie osoby (czy przez jedną osobę w różnych momentach) w taki sposób, iż jego użyciem rządzić będą odmienne standardy semantyczne, mimo że ciągle pozostanie tym samym pojęciem. Zmiany istniejące w poprawnym użyciu danego pojęcia będą różnicami zachowującymi tożsamość danego pojęcia czy znaczenia słowa⁹⁴. Powiemy wówczas, że owa zachowująca tożsamość różnica jest *różnią* (*différance*) tego pojęcia (czy znaczenia)⁹⁵.

⁹¹ Podać można również realistyczną interpretację powyższego zdania. Hare’a krytyka realizmu zaprezentowana została w § 3.3.3.

⁹² Zob. T. Horgan, M. Timmons, *Conceptual Relativity and Metaphysical Realism*, „Philosophical Issues” (12) 2002.

⁹³ **Wariantyzm** M. Timmons i T. Horgan definiują w następujący sposób: „Semantyczne standardy poprawnego zastosowania C1, jeśli C1 zastosowany jest przez osobę P1 w czasie t1, mogą być różne na pewne dozwolone sposoby od semantycznych standardów poprawnego zastosowania pojęcia C2, jeśli C2 zastosowane jest przez osobę P2 w czasie t2, nawet wtedy gdy C1 = C2. Jeśli tak się dzieje, wtenczas zaistniała różnica [*différance*] w pojęciach” lub „V2 Nawet jeśli semantyczne standardy poprawnego zastosowania słowa W, jeśli W zastosowane jest przez osobę P1 w czasie t1, mogą być różne na pewne dozwolone sposoby od semantycznych standardów poprawnego zastosowania W, jeśli W zastosowane jest przez osobę P2 w czasie t2, to znaczenie W użyte przez P1 = znaczenie W użyte przez P2. Jeśli tak się dzieje, wtenczas zaistniała różnica [*différance*] w pojęciach”; tamże, s. 82.

⁹⁴ Zob. tamże, s. 78.

⁹⁵ Por. tamże, s. 79.

Choć, jak uważał Hare⁹⁶, na pierwszy rzut oka może się wydawać, że wariantyzm jest świadectwem na rzecz nondeskrytywizmu, wcale tak być nie musi. Jeśli znaczenia pojęć moralnych można byłoby zdefiniować jako pierwotnie deskrytywne oraz wariantystyczne, Hare straciłby większość swoich mocnych argumentów przeciwko deskrytywizmowi⁹⁷. Jego uzasadnienie dla rozdzielenia znaczenia pojęć i sądów od kryteriów ich użycia oraz krytyka *reductio ad absurdum* deskrytywizmu uderza bowiem nie w deskrytywizm, ale inwariantyzm⁹⁸, który do tej pory był łączony z deskrytywizmem.

§ 4.7 Argument Fregego–Geacha

Jednym z najpoważniejszych zarzutów, jakie stawia się teoriom nondeskrytywistycznym, jest tzw. **argument Fregego–Geacha** (*Frege – Geach Problem*), znany również pod nazwą **problemu uwikłania** (*embedding problem*)⁹⁹. Argument ten wzięł nazwę od autora artykułu *Ascriptivism*¹⁰⁰ – Petera Geacha, który powołując się na Gottloba Fregego¹⁰¹, po raz pierwszy sformułował go jako krytykę podejścia performerów.

⁹⁶ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 92-93.

⁹⁷ Powyżej przedstawioną nondeskrytywistyczną interpretację zmian, które zachodzą w kryteriach użycia pojęć moralnych, krytykował np. J. W. Nickel. Zob. tenże, *Hare's Argument from Linguistic Change* w: „Ethics” (4) 1969, s. 298-302. Zarzuty wobec dystynkcji znaczenia – kryteria użycia można znaleźć w: D. A. Lloyd Thomas, *A Note on Hare's Analysis of „Good”, „Mind”* (288) 1963, s. 562-567.

⁹⁸ **Inwariantyzm** autorzy artykułu definiują w następujący sposób: „Jeśli semantyczne standardy poprawnego użycia pojęcia C1 przez osobę P1 w czasie t1 różnią się od semantycznych standardów użycia pojęcia C2, użytego przez osobę P2 w czasie t2, to C1 ≠ C2”; T. Horgan, M. Timmons, *Conceptual Relativity...*, dz. cyt., s. 77. Podają oni również identyczną definicję nie pojęcia, lecz znaczenia słowa W.

⁹⁹ Zob. P. T. Geach, *Ascriptivism*, „Philosophical Review” (2) 1969, s. 221-225; tenże, *Assertion*, „Philosophical Review” (4) 1965, s. 449-465; J. Searle, *Meaning and Speech Acts*, „Philosophical Review” (4) 1962, s. 423-432; H. N. Castañeda, *Imperatives, Decisions and Oughts*, w: H. N. Castañeda, G. Nakhnikian (ed.), *Morality and the Language of Conduct*, Wayne State University Press, Detroit 1965, s. 219-300.

¹⁰⁰ P. T. Geach, *Ascriptivism...*, dz. cyt.

¹⁰¹ Zob. G. Frege, *Negation*, w: P. Geach, M. Black (ed.), *Translations from the Philosophical Writings of Gottlob Frege*, Oxford University Press, Oxford 1980, s. 129-130.

Istnieje – jak stwierdza J. Searle¹⁰² – warunek adekwatności, który każda analiza znaczenia słowa musi spełniać. Nie udaje się to jednak metodzie proponowanej przez performerów. Wynikiem każdej poprawnej analizy znaczenia słowa (czy morfemu) musi być taki opis jego treści, który byłby poprawny we wszystkich możliwych gramatycznie złożonych rodzajach zdań. Aby teoria metaetyczna mogła być uznana za dobrą, musi ona wyjaśniać całą semantyczną różnorodność dyskursu moralnego, w którym stosowanie podstawowych funktorów logicznych (takich jak negacja, koniunkcja, alternatywa, implikacja itd.) jest praktyką powszednią. Jeśli zgodzimy się na sposób analizy, który proponują performerzy, będzie to jednak oznaczało, iż wnioskowanie:

1. Jeśli A jest złe, to B również jest złe
2. A jest złe

3. B również jest złe

podlega błędowi ekwiwokacji¹⁰³. Aby dedukcja powyższa była poprawna, we wszystkich przesłankach słowo „złe” powinno oznaczać to samo. Jeśli jednak przyjmiemy założenia performerów, to w zdaniu drugim znaczenie „zły” należy traktować jako ewaluatywne („Potępiam (zakazuję) A”), pomimo iż słowo „zły” w zdaniu pierwszym nie może wyrażać preskrypcji, gdyż przesłanka pierwsza nie jest wyrażeniem żadnego potępienia, lecz zdaniem warunkowym. Nie jest przecież tak, że gdy wypowiadamy: „Jeśli robienie jakiejś rzeczy jest złe, namawianie twojego młodszego brata do robienia tego również jest złe”, potępiamy w ten sposób „robienie jakiejś rzeczy” lub że cała nasza wypowiedź jest potępieniem¹⁰⁴. Jeśli zgodzimy się, że moc illokucyjna zdania: „Przyjemność jest dobra” jest zaleceniem, nie znaczy to, że słowo „dobry” nabiera swojego znaczenia za sprawą wykonywanej czynności zalecania, ponieważ moc illokucyjna zdań: „Przyjemność nie jest dobra”, „Nie jest tak, że przyjemność jest dobra”, „Jeśli przyjemność jest dobra, to Bentham miał rację” nie jest zaleceniem. Jak wyraził to Searle:

Twierdzenie: „S używa się, by wykonać czynność C”, które uzyskano, badając proste zdania oznajmujące w czasie teraźniejszym, zawierające S, nie wyja-

¹⁰² Zob. J. Searle, *Czynności...*, dz. cyt., s. 175.

¹⁰³ Por. P. T. Geach, *Ascriptivism...*, dz. cyt., s. 223.

¹⁰⁴ Por. tenże, *Assertion...*, dz. cyt., s. 463.

śnia występowania S w wielu rodzajach zdań nie będących prostymi zdaniami oznajmującymi w czasie teraźniejszym¹⁰⁵.

W taki sposób argumentował również H. N. Castañeda, twierdząc, że oceny różnią się od imperatywów tym, że te drugie używane mogą być jedynie asertywnie, zatem nie mogą być stosowane w kontekstach złożonych¹⁰⁶.

Zdaniem P. Geacha, przedstawiciele **askryptywizmu**¹⁰⁷ nie mogą sobie poradzić z poprawną analizą złożonych kontekstów gramatycznych, ignorują oni bowiem uczynione przez Fregego bardzo ważne rozróżnienie pomiędzy predykacją a asercją, które w języku moralnym koresponduje z rozróżnieniem na predykację i np. nakazywanie. Czym innym bowiem jest nakazać „P”, co wymaga od nas podpisania się pod (asercji) „P”, czym innym natomiast orzec o czymś „P”. Według Geacha, aby wyjaśnić znaczenie słowa „zły”, musimy uczynić to poprzez czynność orzekania „zła” o danej rzeczy, takie orzekanie natomiast może być dokonane bez jakiegokolwiek aktu potępienia czy asercji¹⁰⁸.

Próbując odpowiedzieć na powyższy zarzut, Hare w artykule *Meaning and Speech Acts* postawił trzy tezy¹⁰⁹:

I. „Dobry” i „jest” posiadają to samo znaczenie w zdaniach prostych i złożonych.

II. Wyjaśnienie ich znaczenia musi obejmować fakt, że standardowo są one używane do wykonywania czynności nakazywania („dobry”) oraz stwierdzania („jest”).

III. W zdaniach warunkowych owe czynności mowy nie są wykonywane.

Według Hare’a, akt obiecywania nie jest wykonywany przez osobę, która wypowiada pytanie: „Czy ja obiecuję?”, pomimo iż znaczenie wyrazu obiecuję, powinniśmy wyjaśnić tak, jak wskazuje na to

¹⁰⁵ J. Searle, *Czynności...*, dz. cyt., s. 176-177.

¹⁰⁶ Zob. H. N. Castañeda, *Imperatives, Oughts, and Moral Oughts*, „The Australasian Journal of Philosophy” (3) 1966, s. 283.

¹⁰⁷ **Askryptywizm** jest terminem szerszym niż ekspresywizm. Jego znaczenie pokrywa się ze znaczeniem pojęcia nondeskryptywizm, jeśli nondeskryptywizm będziemy pojmowali szeroko, tj. jako sposób analizy znaczeń nie tylko w etyce, ale również w epistemologii czy teorii komunikacji. Oprócz emotywizmu czy preskrytywizmu dla P. T. Geacha przykładem takiej teorii jest również np. performatywna teoria prawdy Strawsona. Zob. P. T. Geach, *Ascriptivism...*, dz. cyt.

¹⁰⁸ Zob. tenże, *Ascriptivism...*, dz. cyt., s. 223-224.

¹⁰⁹ Por. R. M. Hare, *Meaning...*, dz. cyt., s. 93.

punkt II. Nie zmienia to jednak faktu, że to słowo we wszystkich kontekstach znaczy to samo (I). Aby zrozumieć powyższe pytanie, musimy wiedzieć, że pytamy o obiecywanie – czynność illokucyjną, którą standardowo wykonujemy, gdy mówimy: „Obiecuję, że...”. Ktoś, kto nie rozumie, iż słowo „obiecuję” służy do dokonywania takiego aktu, nie rozumie również powyższego pytania. Podobna zależność zachodzi, gdy mówimy o negacji. Znamy znaczenie zaprzeczenia „Ja nie obiecuję” tylko wtedy, gdy wiemy, co oznacza zdanie „Ja obiecuję”, oraz znamy zasady użycia negowania¹¹⁰. „Dlaczego nie moglibyśmy wyjaśnić znaczenia zdania «To jest dobry film» częściowo w terminach czynności chwaleń tego filmu, a później wyjaśnić znaczenie «Czy to jest dobry film?» i «To nie jest dobry film» pośrednio, jako transformację znaczenia zdania pierwszego?”¹¹¹ Analogicznie, aby poznać znaczenie „Jeśli kot jest na macie, to mruczy”, musimy wiedzieć (1), jakie znaczenie ma warunkowa forma zdaniowa, co rozpoznamy, jeśli potrafimy używać *modus ponens*¹¹², oraz (2) jakie znaczenie posiadałyby jej zdania składowe, gdyby nie były związane przez powyższą formę zdaniową. Odpowiedź na (2) znamy, jeśli wiemy, że te zdania, gdy nie są częścią zdania warunkowego, używane są do stwierdzania, oraz jeśli wiemy, do stwierdzania czego są one używane¹¹³.

§ 4.8 Hare’a teoria czynności mowy

Aby lepiej zrozumieć powyższą odpowiedź Hare’a na *problem uwiłkiania*, należy przyrzeć się przedstawionej przez niego autorskiej teorii czynności mowy¹¹⁴. Według niego, aby wykonać jakikolwiek akt mowy, zdanie musi zawierać tropik, frastik, neustik i klistik. **Tropik** (*mood sign, tropic*) wyznacza, jaki rodzaj czynności mowy jest wykonywany (imperatyw, stwierdzenie faktu, pytanie).¹¹⁵ **Fra-**

¹¹⁰ Zob. tamże, s. 82, 84.

¹¹¹ Zob. tamże, s. 85.

¹¹² Zob. tamże, s. 86.

¹¹³ Zob. tamże, s. 87, 88.

¹¹⁴ Swoją teorię czynności mowy Hare wstępnie przedstawił już w *The Language of Morals*. Jednak znaczenia takich pojęć, jak frastik czy neustik, od tego czasu uległy przeobrażeniu. Najpełniejszą i dojrzałą formę jego teoria czynności mowy przybrała w pracy: *Some Subatomic Particles of Logic*, w: tenże, *Objective...*, dz. cyt., s. 28-42.

¹¹⁵ Zob. tenże, *Meaning...*, dz. cyt., s. 91.

stik (*phrastic*) jest odpowiednikiem treści propozycjonalnej, jest nazwą tego, co jest stwierdzane, o co pytamy lub co nakazujemy. Jest on takim układem słów, że dodając do niego tropik nadajemy mu konkretny tryb – rodzaj mocy illokucyjnej, a dodając neustik, możemy podpisać się pod tym, co zostało w nim powiedziane¹¹⁶. **Neustik** jest cechą zdania, która pokazuje, czy nadawca podpisuje się pod tym, co jest mówione – jest wyrazem zaangażowania w akt mowy. W języku pisanym zazwyczaj jego występowanie komunikowane jest przez nieużywanie cudzysłowia. **Klistik** natomiast jest znakiem, który wskazuje, że to, co zdanie mówi, jest wszystkim, co jest w nim wypowiedziane. Jest to znak kompletności aktu mowy¹¹⁷. Funkcję tę w języku pisanym najczęściej pełni kropka. Zdaniem Hare’a, nie są to wszystkie, jak je nazywał – subatomowe części logiki, był on raczej skłonny twierdzić, że jest ich więcej. Cztery wymienione powyżej wydają się jednak najważniejsze¹¹⁸.

Zasadnicza zmiana, jaką wniósł Hare do klasycznej teorii aktów mowy Austina, nie polegała wyłącznie na wprowadzeniu klistika, lecz przede wszystkim na rozdzieleniu neustika od tropika, które standardowo są ze sobą łączone w pojęciu mocy illokucyjnej zdania. Według Hare’a, trudność związana z argumentem Fregego-Geacha przestałaby istnieć, gdyby ta dychotomia była powszechniej znana. Powyższy podział mocy illokucyjnej usuwa bowiem problematyczność stwierdzenia, że tropik (rodzaj czynności: obiecywanie, proszenie, nakazywanie, pytanie itd.) nie może być częścią zdania podrzędnego, ponieważ nikt się pod nim nie podpisuje. Według Hare’a, zdania podrzędne zawierają w sobie tropik, choć nie mogą zawierać neustika¹¹⁹. Jedynie neustik nie może występować w zdaniu podrzędnym¹²⁰.

Jeśli neustik oznaczymy przez „|”, tropik przez „Δ” (Δ₁ – stwierdzenie; Δ₂ – komenda; Δ₃ – pytanie; Δ₄ – przeczenie; Δ₅ – powinność), frastik będzie zbiorem słów, a funkcję klistika pełnić będzie znak kropki – „.”, problematyczne zdania można przedstawić w następujący symboliczny sposób:

¹¹⁶ Zob. tenże, *Some Subatomic...*, dz. cyt., s. 40.

¹¹⁷ Zob. tamże oraz tenże, *Imperatives, Prescriptions...*, dz. cyt., s. 45.

¹¹⁸ Zob. tenże, *Some Subatomic...*, dz. cyt., s. 40.

¹¹⁹ Zob. tamże, s. 41.

¹²⁰ Zob. tenże, *Meaning...*, dz. cyt., s. 91.

(1) Kot jest na macie.	⊢Δ ₁ Kot jest na macie.
(2) Jeśli kot jest na macie, to mruczy.	⊢ (Δ ₁ Kot jest na macie → Δ ₁ Kot mruczy).
(3) Jeśli kot jest na macie, nakam go.	⊢ (Δ ₁ Kot jest na macie → Δ ₂ Nakam go).
(4) Czy kot jest na macie?	⊢Δ ₃ Kot jest na macie.
(5) Wypierz matę albo wyczyść kota.	⊢ (Δ ₄ Wypierz matę ∨ Δ ₄ Wyczyść kota).
(6) Kot nie jest na macie.	⊢Δ ₄ Kot jest na macie.
(7) Kot powinien być na macie.	⊢Δ ₅ Kot jest na macie.

Tabela 5. Słowa moralne w kontekstach złożonych

Łączenie neustika z tropikiem powodowane było m.in. tym, że niektóre akty mowy, aby mogły być dokonane poprawnie, wymagają ich obydwu. Np. pomimo iż zdania (2) i (3) posiadają tropik odpowiedzialny za tryb oznajmujący, bez dodania do nich neustika nie mogą być one stwierdzeniami¹²¹. W jaki sposób rozpoznać jednak, czy w zdaniu jest zawarty neustik? Można postawić zarzut, że jeśli nie istnieje jakiś jego wskaźnik, to powyższy pomysł jest mnożeniem bytów ponad potrzebę. Hare jednak wierzył, że wskaźniki takie istnieją, a z pewnością mogłyby istnieć na mocy konwencjonalnej decyzji, ponieważ są potrzebne. W artykule *Imperative Sentences*, analizując fragmenty *Principia Mathematica*, udowodnił, że użyty tam znak asercji („|”) pełni trzy funkcje, z których każda odpowiada jednemu z subatomowych części jego logiki: (a) funkcję znaku kompletności – *klistik*; (b) znaku użycia – *neustik*; oraz (c) znaku użytego trybu (stwierdzenia) – *tropik*¹²². Według Hare’a, istnieją dwa główne powody, dla których znak asercji przestał być powszechnie stosowany: po pierwsze, stało się tak za sprawą pierwszego Wittgensteina, który analizując zdania poza ich użyciem, uznał znak asercji za niepotrzebny¹²³; po drugie, logicy zastąpili znak zaangażowania brakiem cudzysłowu, który metodycznie stosowany, dobrze spełnia

¹²¹ Zob. tamże.

¹²² Zob. tenże, *Imperative Sentences...*, dz. cyt., s. 23.

¹²³ Zob. L. Wittgenstein, *Tractatus logico-philosophicus*, tłum. B. Wolniewicz, Wydawnictwo Naukowe PWN, Warszawa 2000, t. 4, 442, s. 37.

swoją rolę¹²⁴. Jeśli – twierdził – możemy udowodnić, że istnieje zbiór wyrazów, które wskazują na zaangażowanie lub jego brak (wskaźniki neustika), a zarazem nie są wskaźnikami tropika, to jego teoria subatomowych części logiki będzie mieć mocne podstawy. Słowami takimi są m.in.: sugerować i nalegać. Innym czytelnym wskaźnikiem neustika w języku pisanym jest nieobecność cudzołowu¹²⁵. Konwencjonalnym wyrazem zaangażowania jest również podpisanie się na czeku czy złożenie przysięgi w sądzie¹²⁶.

§ 4.9 Konwencje a intencje

Zamierzana lub interpretowana moc illokucyjna wyrażenia, podobnie jak jego zamierzone lub interpretowane znaczenie, nie muszą być identyczne z rzeczywistością, tj. standardową mocą i rzeczywistym, tj. standardowym znaczeniem tego, co się mówi¹²⁷. Zdaniem Hare’a, nie można identyfikować mocy illokucyjnej danej wypowiedzi z tym, jak traktuje tę wypowiedź jej konkretny użytkownik¹²⁸. Nie możemy mylić znaczenia zdania, ze znaczeniem, które nadaje wypowiedzi nadawca (*sentence-meaning* z *utterer’s meaning*), choć są one ściśle połączone. Znaczenie pełnego zdania wypowiedzianego *explicite* determinuje jego treść. Jeśli ktoś mówi „moja lewa ręka”, musi odnosić się do swojej lewej ręki albo źle używa słów, które wypowiada. Osoba, która nie stosuje się do logiki danej wypowiedzi, nie zna znaczeń zawartych w niej terminów lub źle używa słów. Z drugiej jednak strony, powinniśmy pamiętać, iż nie istnieją sztywne wskaźniki mocy illokucyjnej – wyrazy, które jednoznacznie determinują moc illokucyjną danej wypowiedzi¹²⁹. Słowa typowo moralne mogą być bowiem użyte w znaczeniu deskryptywnym lub ewaluatywnym¹³⁰.

¹²⁴ Zob. R. M. Hare, *Imperative Sentences...*, dz. cyt., s. 24.

¹²⁵ Zob. tenże, *Some Subatomic...*, dz. cyt., s. 30-31.

¹²⁶ Inną problematyczną kwestią jest to, czy gdyby neustik został wypowiedziany, byłby znakiem wystarczającym, aby wyrażać zaangażowanie nadawcy. Zreferowanie dyskusji na ten temat, toczącej się pomiędzy R. M. Hare’em a D. Davidsonem, znaleźć można w: M. S. Green, *On The Autonomy of Linguistic Meaning*, „Mind”, (422) 1997, s. 217-243.

¹²⁷ Zob. J. Szymura, *Język...*, dz. cyt., s. 266.

¹²⁸ Zob. R. M. Hare, *Appendix. Reply to...*, dz. cyt., s. 96.

¹²⁹ Zob. tenże, *The Language...*, dz. cyt., s. 180.

¹³⁰ Zob. tenże, *Myślenie...*, dz. cyt., s. 36.

Aby się dowiedzieć, co znaczy słowo „powinno się”, Hare nie pyta, w jakich celach ludzie świadomie używają tego słowa, co chcą przez nie osiągnąć, nie interesuje go perlokucyjny aspekt wypowiedzi ani intencja jej nadawcy, lecz to, jakie są warunki i implikacje uprawnionego zastosowania tych słów. Owymi warunkami wygłaszania niewadliwych twierdzeń jest używanie ich zgodnie z ustalonymi prawami logiki danej gry językowej. Preferowany przez niego sposób analizy polega na opisie nie intencji użytkowników języka, ale konwencjonalnych reguł jego użycia. Wydaje się, iż w tej kwestii jego stanowisko pokrywało się z przekonaniem J. L. Austina. Według Austina, a wbrew P. Strawsonowi, akt illokucyjny jest aktem, który wykonuje się zgodnie z konwencją; jest on jednak konwencjonalny nie w tym sensie, że może być całkowicie oparty na regułach języka, lecz dlatego, że jest aktem, który wykonuje się zgodnie z ustalonymi regułami¹³¹. Hare nie był więc zwolennikiem *theory of communication-intention*. Uchroniło go to przed utożsamieniem perlokucji z illokucją. Zwolennicy tej koncepcji uznają bowiem, że konwencje językowe należy wyjaśnić w kategoriach intencji mówiących i słuchaczy, gdyż są one genetycznie i eksplanacyjnie bardziej pierwotne. Hare jednak uważał, tak jak Austin i Searle¹³², że konwencje języka i mowy (logika) są bardziej podstawowe niż intencje i za ich pomocą wyjaśnić można znaczenia i funkcje, jakie użytkownicy języka intencjonalnie wiążą z danymi wyrażeniami. Znaczenie ewaluatywne nie jest jedynie wyrazem intencji wykonania danej czynności illokucyjnej, lecz jest cechą konwencjonalną, którą opisać można podając reguły logiki wypowiedzi posiadających to znaczenie. Wśród filozofów akceptujących powyższy pogląd istnieją jednak rozbieżności w pojmowaniu owych konwencji. Spór między R. M. Hare’em a J. Searle’em, toczący się w kwestii przejścia od „jest” do „powinien”, który przedstawimy w paragrafie 4.12, dotyczy właśnie charakteru owych reguł.

§ 4.10 Logika moralnej gry językowej a znaczenie

Hare był zdania, podobnie jak Austin, że oprócz słów, które są nazwami i posiadają odniesienie, można wyszczególnić także te, które pełnią inne funkcje i nie posiadają referencji. Austin nazwał je „sło-

¹³¹ Zob. J. Szymura, *Język...*, dz. cyt., s. 265.

¹³² Por. tamże, s. 258.

wami adaptacyjnymi” (*adjuster-words*). Wyrażenia te nie posiadają sensu empirycznego, mają charakter „jałowo-formalny” (*substantive-hungry*) i można je scharakteryzować wyłącznie przez podanie funkcji, jaką pełnią w zdaniu¹³³. Można powiedzieć, że słowa adaptacyjne mają charakter przymiotników „modyfikujących”; nie zaś „atrybutywnych czy determinujących”¹³⁴. Niedostrzeżenie faktu, iż *adjuster-words* są jedynie z pozoru nazwami i brak im sensu empirycznego, jest przyczyną najpoważniejszych błędów w filozofii. Traktowanie ich jako deskrypcji prowadzi do poszukiwania domniemyanych bytów, które słowa te mają rzekomo opisywać¹³⁵. Słowa adaptacyjnymi oprócz takich, jak „prawdziwy”, „rzeczywisty”, są także takie, jak „dobry” i „zły”. W konsekwencji więc, zdaniem Austina, poszukiwanie definicji „dobra” i „zła” jest poszukiwaniem sensu słów, które go nie posiadają – filozof może jedynie opisać ich funkcję.

Hare nie był tak radykalny w swoich poglądach. Jego zdaniem, choć pojęcia moralne nie są nazwami, to posiadają znaczenie. Nie jest ono jednak wyznaczone, jak głosili emotywiści, przez ekspresję emocji czy sposób perlokucyjnego oddziaływania. Istnieją wyrażenia, które nie mają desygnatów i nie służą do wywoływania emocji czy wywoływania postaw, a ich znaczenie znamy bardzo dobrze. Takimi słowami są: „każdy”, „niektóre”, „jeżeli...”, „to” itd. Podstawowe słowa moralne mają więc charakter operatorów logicznych. Pomimo iż Hare nie stosował takiego terminu, wydaje się, że ma on w tym względzie rację, według niego bowiem słowa moralne wprowadzają pewne relacje, które można opisać formalnie. Tworzą one logikę moralności i podobnie jak funktory logiczne nie mają znaczeń semantycznych, opisują jedynie pewną strukturę dla ewentualnej semantyki. Znaczenie terminów etycznych jest definiowane poprzez funkcję, którą pełnią w zdaniu. Wskazuje ona, że słowa czysto ewaluatywne, takie jak „powinien”, „musi”, „dobry”, mają znaczenie wyłącznie formalne i są w języku naturalnym odpowiednikami operatorów logiki deontycznej¹³⁶. Jest to jedna z najbardziej fundamentalnych jego tez. Aby opisać znaczenie powyższych słów należy

¹³³ Por. tamże, s. 97. Zob. J. L. Austin, *Mówienie...*, dz. cyt., s. 101-102.

¹³⁴ Por. J. Szymura, *Język...*, dz. cyt., s. 99.

¹³⁵ Por. tamże, s. 101-102.

¹³⁶ Zob. R. M. Hare, *Ontology...*, dz. cyt., s. 98. Zob. również tenże, *Sorting...*, dz. cyt., s. 6.

zaprezentować ich własności logiczne¹³⁷, i z tego powodu zadaniem etyka powinno być studiowanie logiki języka moralnego¹³⁸. **Logiką moralności** nazywał Hare zbiór reguł rządzących moralną grą językową, które mają charakter pojęciowej konieczności¹³⁹ i nie odnoszą się do reguł syntaktycznych, lecz semantyki i pragmatyki. „Chude pojęcia moralne” nie są funktorami prawdziwościowymi. Ich funkcją jest dodawanie do znaczenia deskryptywnego znaczenia ewaluatywnego, które w konsekwencji zmienia sens całej wypowiedzi. Są one modyfikatorami gry językowej – z opisowej na oceniającą. W języku czynności mowy powiedzieć można, iż zmieniają moc illokucyjną wypowiedzi, modyfikując zarazem jej cechy semantyczne i pragmatyczne. „Chude pojęcia moralne” są podobne do funktorów logicznych. Aby to uzasadnić, w *Mysleniu moralnym*¹⁴⁰ Hare stwierdził za W. Sellarsem, iż o znaczeniach takich słów, jak „powinien” i „musi”, jesteśmy w stanie dowiedzieć się wszystkiego, jeśli wiemy, jakie wnioski możemy wyciągnąć ze zdań, które je zawierają¹⁴¹. Jak pisał Sellars: „Znaczenie terminu logicznego jest funkcją jego udziału w implikacjach zdań, w jakich ono występuje. To, że powyższy fakt przynajmniej częściowo odnosi się do znaczenia *każdego* wyrazu, jest pewnie już mniej oczywiste, ale, rzecz jasna, równie ważne”¹⁴².

Hare w żadnej ze swych prac nie podał skończonej, sprawozdawczej definicji głównych pojęć moralnych. Uważał bowiem, iż uniemożliwia nam to bogactwo sposobów użycia języka potocznego¹⁴³. Z drugiej jednak strony opisywał ich własności, przedstawiał relacje, które między nimi zachodzą, oraz próbował skonstruować ich formalne odpowiedniki – takie definicje projektujące, które byłyby najbardziej adekwatnymi ekwiwalentami pojęć języka potocznego w standardowym ich użyciu¹⁴⁴. W tym celu Hare starał się zdefiniować takie wyrazy, jak „dobry”, „zły”, „słuszny”, „niesłuszny” poprzez słowo „powinno się”. Z tego względu należy uznać je za pierwotny termin etyczny.

¹³⁷ Zob. tenże, *Supervenience...*, dz. cyt., s. 75-80.

¹³⁸ Zob. tenże, *Myslenie...*, dz. cyt., s. 13-14.

¹³⁹ Tenże, *Sorting out Ethics*, s. 7.

¹⁴⁰ Zob. tenże, *Myslenie...*, dz. cyt., s. 13.

¹⁴¹ W. Sellars, *On Reasoning about Values*, „American Philosophical Quarterly” (17) 1980, s. 81.

¹⁴² Tamże.

¹⁴³ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 180.

¹⁴⁴ Zob. tamże, s. 3.

Słuszny/Właściwy¹⁴⁵

1. Pierwsze znaczenie:
 - a) Nie jest słusznym czynić A = X nie powinien czynić A
 - b) Nie było słuszne uczynienie A = X nie powinien być uczynić A
 - c) Nie byłoby słuszne uczynić A = Jeśli X uczyniłby A, zrobiłby to, czego nie powinien był robić
2. Drugie znaczenie:
 - a) Słusznie, że X zrobił A = X robiąc A, zrobił to, co powinien
 - b) Nie jest niesłuszne, że X zrobił A = X robiąc A, nie zrobił nic, czego nie powinien
3. Trzecie znaczenie:
 - a) Właściwe A = A które powinno być (powinno było być) wybrane
 - b) Nie zadzwonił do właściwego domu = Nie zadzwonił do domu, który powinien był wybrać, aby zadzwonić

Lepszy niż¹⁴⁶

A jest lepszym X niż B = Jeśli ktoś wybierając X, wybiera B, to powinien wybrać A¹⁴⁷.

Według Hare'a, podanie definicji słowa „powinno się” polega na przedstawieniu formalnych własności zdań, w których ono występuje¹⁴⁸, implikacji, które zdania te pociągają. Zatem aby rozpoznać sądy moralne spośród zbioru innych wypowiedzi, musimy wziąć pod uwagę ich preskryptywność, uniwersalizowalność i nadrzędność¹⁴⁹.

Powinno się¹⁵⁰

Znaczenie wyrażenia „powinno się” jest takie, iż sądy, które je zawierają, są uniwersalizowalnymi zaleceniami o charakterze nadrzędności.

¹⁴⁵ Por. tamże, s. 182.

¹⁴⁶ Ponieważ zdaniem Hare'a o wiele prościej jest podać definicję „dobry” w gramatycznym stopniu wyższym („lepszy niż”), nie przedstawił on definicji słowa dobry w swoim brzmieniu podstawowym. Krytykę powyższej definicji znaleźć można w: M. K. Rennie, *On Hare's „Better”*, „Noûs” (1) 1968, s. 75-79.

¹⁴⁷ R. M. Hare, *The Language...*, dz. cyt., s. 184.

¹⁴⁸ Zob. tenże, *Moral Philosophy*, wywiad z R. M. Hare'em przeprowadzony przez B. Magee w programie „Men of Ideas”, BBC, London 1978. Pobrane z Internetu dnia 12.06.2004 z <<http://groups.yahoo.com/group/analytic/message/1608>>.

¹⁴⁹ Zob. tenże, *Myślenie...*, dz. cyt., s. 72-74.

¹⁵⁰ Por. H. J. Gensler, *The Prescriptivism Incompleteness Theorem*, „Mind” (340) 1976, s. 589.

§ 4.11 Logika imperatywów

Aby udowodnić, że wszystkie zalecenia, a zatem i sądy moralne, podlegają prawom logiki, Hare starał się wykazać, że nawet najpopolitsza i najbardziej problematyczna z punktu widzenia logiki klasycznej forma zaleceń, jaką są imperatywy, podlega jej prawom. Wierzył bowiem, że gdyby udało się dowieść, iż imperatywy, które w swojej warstwie powierzchniowej, składni i gramatyce tak zasadniczo różniące się od przekonań, wchodzą w relacje inferencyjne, nie istniałyby już żadne ważne przeszkody, które uniemożliwiałyby przyjęcie istnienia logiki sądów moralnych traktowanych jako zalecenia¹⁵¹. Byłby to dowód, że rozpoczęta przez emotywistów analiza w duchu nondeskryptywizmu może być wzbogacona o racjonalizm.

Niektórzy uważają jednak, że logika imperatywów nie istnieje. Twierdził tak np. Karl Popper¹⁵². Również Alf Ross udowodniał, że wnioskowanie o formie: „Jeśli idziesz, weź swój płaszcz; Idź; Zatem weź swój płaszcz” jest niepoprawne¹⁵³. Zdaniem P. Geacha¹⁵⁴, wypowiedź, która nie może być ani prawdziwa, ani fałszywa, nie może również być ani przesłanką, ani wnioskiem jakiegokolwiek rozumowania. Jeśli imperatywy nie mają wartości logicznych, a nie ma co do tego wątpliwości, to racjonalna argumentacja w etyce nie miałaby miejsca. Polemiki etyczne opierają się jednak o zasady racjonalności, zatem sądy etyczne nie mogą być podobne do imperatywów. Taki sposób argumentowania był jedną z głównych metod ataku na uniwersalny preskrytywizm. Jak pisaliśmy w paragrafie 4.1.2, przeciwko teorii Hare'a użył go Jonathan Harrison¹⁵⁵.

Hare próbował uzasadnić możliwości istnienia logiki imperatywów już w *The Language of Morals*. Starał się wykazać, iż imperatywy mogą być wzajemnie niezgodne (*inconsistent*), dzięki czemu nazwać je można sprzecznymi (*self-contradictory*). Dzieje się tak np. w wypadku wypowiedzi „Zamknij drzwi!” i „Nie zamykaj drzwi!”¹⁵⁶,

¹⁵¹ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 2.

¹⁵² Zob. K. Popper, *What Can Logic Do for Philosophy*, „Proceedings of the Aristotelian Society” (22) 1948, s. 154.

¹⁵³ Zob. A. Ross, *Imperatives and Logic*, „Philosophy of Science” (11) 1944, s. 30-46.

¹⁵⁴ Por. P. T. Geach, *Assertion...*, dz. cyt., s. 449-465.

¹⁵⁵ Por. przypis 16.

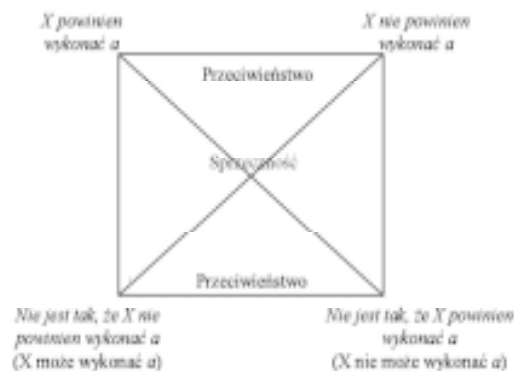
¹⁵⁶ Zob. R. M. Hare, *Imperatives, Prescriptions...*, dz. cyt., s. 48.

które nie mogą być wykonane jednocześnie (nikt nie może zarazem zamknąć i nie zamknąć drzwi). Generalna definicja niezgodności brzmi następująco: „jeśli dwa nakazy nie mogą być spełnione na mocy logiki (tj. jeśli stwierdzenia mówiące, że są one spełnione, są zarazem niezgodne), to nakazy te są również niezgodne¹⁵⁷.” Według Hare’a imperatywy obowiązują więc następujące prawo:

Prawo niesprzeczności imperatywów

$$\sim [!(a \text{ robi } A)] \& ! \sim (a \text{ robi } A)]$$

Zdaniem sprzecznym do „Smith powinien wykonać *a*” nie jest „Smith nie powinien wykonać *a*”, ale „Nie jest tak, że Smith powinien wykonać *a*”, natomiast negacją „Smith nie powinien wykonać *a*” jest „Nie jest tak, że Smith nie powinien wykonać *a*.” „Powinno się” i „nie powinno się” nie są więc wyrażeniami sprzecznymi, ale przeciwstawnymi¹⁵⁸. Natomiast zdanie „*X* może wykonać *a*” znaczy „Nie jest tak, że *X* nie powinien wykonać *a*”¹⁵⁹. Powyższe relacje można przedstawić za pomocą kwadratu logicznego:



Rysunek 5. Kwadrat logiczny powinności

¹⁵⁷ Tenże, *Practical Inferences*, w: tenże, *Practical Inferences...*, dz. cyt., s. 71.

¹⁵⁸ Por. tenże, *Some Alleged Differences between Imperatives and Indicatives*, w: tenże, *Practical Inferences...*, dz. cyt., s. 35. Zarzut, iż Hare myli powinność z byciem dozwolonym można znaleźć w: D. P. Gauthier, *Hare's Debtors*, „Mind” (307) 1968, s. 400-405. Odpowiedzi na powyższy zarzut udzielił M. T. Thornton w: tenże, *Hare's View of Morality*, „Mind” (320) 1971, s. 617-619.

¹⁵⁹ Por. R. M. Hare, *Some Alleged Differences...*, dz. cyt., s. 36.

Pogląd, że imperatywy podlegają również innym regułom logicznym, Hare uzasadniał powołując się na powyższe prawo sprzeczności imperatywów. Jak twierdził:

Jeżeli możemy w ten sposób pokazać, że jakieś imperatywy mogą zostać nazwane sprzecznymi, możemy rozpocząć tworzenie logiki imperatywów i preskrypcji. [...] Jeżeli mogą występować sprzeczności pomiędzy imperatywnymi aktami mowy, będzie istnieć przynajmniej jeden możliwy rodzaj wnioskowania. Np. jeśli mówienie „Zamknij drzwi lub nie zamykaj drzwi” jest wewnętrznie sprzeczne, możemy, w pewnym sensie, „wynioskować” z „Zamknij drzwi” i znaczenia negacji: „Nie nie zamykaj drzwi” („Nie zostawiaj ich otwartych”)¹⁶⁰.

Ponieważ nie chcemy, aby nasze wypowiedzi były wewnętrznie sprzeczne, a są one często ze sobą treściowo połączone, nie możemy dopuszczać, żeby wygłaszane przez nas koniunkcje imperatywów były niezgodne, skoro wiemy, że mogą takie być. W rezultacie wydając jakiś nakaz, jeśli nie chcemy być wewnętrznie sprzeczni, musimy kierować się prawami logiki¹⁶¹. Logika imperatywów mówi nam bowiem, co zrobić, aby akceptowany przez nas zbiór imperatywów (oraz stwierdzeń) nie był wewnętrznie sprzeczny¹⁶².

Jeśli przyjrzymy się różnym użyciom zdań, których elementami składowymi są imperatywy, zauważymy, że zawierają one również takie słowa, jak „wszystkie”, „i”, „lub”, „niektóre”, „jeżeli...to” itd. Takimi zdaniami są np.: „Przeczytaj tę książkę lub napisz wypracowanie”; „Jeśli przeczytasz tę książkę, napisz wypracowanie”. Jeżeli w języku potocznym funkcje logiczne są poprawnie stosowane w zdaniach z imperatywami, według Hare’a jest to kolejny argument za tym, że istnieje logika zaleceń¹⁶³. Co więcej, można wskazać również prawidłowe rozumowania, które zawierają imperatywy jako przesłanki. Najbardziej popularną ich formą jest sylogizm praktyczny, np.:

Jeśli dojedziesz do skrzyżowania, skręć w prawo.

I Jeśli skręcisz w prawo, daj odpowiedni sygnał.

Jeśli dojedziesz do skrzyżowania, daj odpowiedni sygnał.

Istnieją również poprawne wnioskowania, które składają się z przesłanek o różnym tropiku. Można wydedukować jednost-

¹⁶⁰ Tenże, *Imperatives, Prescriptions...*, dz. cyt., s. 45.

¹⁶¹ Por. R. M. Hare, *Practical...*, dz. cyt., s. 71.

¹⁶² Por. tamże, s. 71.

¹⁶³ Zob. tenże, *The Language...*, dz. cyt., s. 25-26.

kową preskrypcję z koniunkcji sądu opisowego i zalecenia, np.¹⁶⁴:

Idź do największego sklepu spożywczego w Oksfordzie.

II Grimly Hughes jest największym sklepem spożywczym w Oksfordzie.

Idź do *Grimly Hughes*.

Kolejnym przykładem jest¹⁶⁵:

Zanieś wszystkie pudełka na stację.

III To jest jedno z pudełek.

Zanieś pudełko na stację.

O istnieniu logiki imperatywów przekonują nas również przykłady, które jawią nam się jako ewidentnie niepoprawne¹⁶⁶:

Zanieś wszystkie pudełka na stację.

IV To jest jedno z pudełek.

Nie zanoś tego pudełka.

Powyższe relacje logiczne pomiędzy imperatywami a oznajmieniami są możliwe, ponieważ istnieje co najmniej jeden element imperatywnych aktów mowy, który, będąc częścią wspólną zarówno imperatywów, jak i twierdzeń, jest podmiotem własności logicznych. Jest nim frastik¹⁶⁷. Z tego powodu każda reguła logiki formalnej, która znana jest jako prawo zdań oznajmujących, ma swoją moc również w odniesieniu do imperatywów¹⁶⁸.

Pod wpływem przedstawionych powyżej argumentów wielu autorów zaaprobowano stanowisko Hare'a. Jak pisze A. MacIntyre:

Hare był w istocie pionierem w badaniach logicznych nad imperatywami. Wykazał, że w dyskursie imperatywnym wnioski mogą wynikać z przesłanek

¹⁶⁴ Por. tamże, s. 35

¹⁶⁵ Tamże, s. 27.

¹⁶⁶ Tamże, s. 25.

¹⁶⁷ Zob. tenże, *Imperatives, Prescriptions...*, dz. cyt., s. 45; tenże, *Imperative Sentences...*, dz. cyt., s. 15.

¹⁶⁸ Zob. tamże, s. 15, 16.

w zupełnie oczywisty sposób, nie łamiąc przy tym zwykłych reguł wynika-
nia¹⁶⁹.

Wielu filozofów rozwijało ten projekt. Zdaniem H. N. Castañedy, inferencje z użyciem imperatywów są w pełni poprawne¹⁷⁰; również Lars Berström bronił logiki imperatywów¹⁷¹. Jednak kilka powyższych przykładów intuicyjnie poprawnych i niepoprawnych dedukcji z pewnością nie może stanowić wyjaśnienia tego skomplikowanego problemu. Istnieją dwie główne przeszkody na drodze do akceptacji stanowiska Hare'a. Po pierwsze, można podważyć tezę, że logika imperatywów jest identyczna z logiką zdań orzekających, prezentując wnioski z przesłankami imperatywnymi, które, według praw tej drugiej, jawią się jako paradoksalne. Po drugie, krytykowano niejasność owych logicznych relacji. Jeśli nawet prawa obowiązujące wśród imperatywów zostaną ustalone oraz przychylimy się do wniosku, że są one izomorficzne z klasycznym rachunkiem zdań, pozostanie otwarte pytanie, jaki mają one charakter, skoro imperatywy nie są ani prawdziwe, ani fałszywe. Czym są wartości semantyczne komend? Co stanowi kryterium poprawności wniosku z użyciem imperatywów¹⁷²?

Jedną z pierwszych odpowiedzi na powyższe pytania była teoria „logiki spełniania” (*logic of satisfaction*¹⁷³). Zamiast pojęcia prawdy czy fałszu, stosuje się w tym przypadku pojęcie spełnienia (*fulfilled, satisfied*) lub niespełnienia nakazu. Imperatyw traktuje się jako spełniony jeśli sytuacja czy czyn, który on nakazuje, stała się faktem, tj. jeśli korespondujące z nim twierdzenie (np. „Janie, zrób A!” – „Jan robi A”) jest prawdziwe. Aby się dowiedzieć, czy dane wnioskiwane jest poprawne, musimy zmienić tropik przesłanek i konkluzji z imperatywnego na stwierdzający, a następnie zbadać, czy tak zmie-

¹⁶⁹ A. MacIntyre, *Krótko...*, dz. cyt., s. 325.

¹⁷⁰ Zob. H. N. Castañeda, *Imperatives, Oughts...*, dz. cyt., s. 283; tenże, *Imperative Reasonings*, „Philosophy and Phenomenological Research” (1) 1960, s. 21.

¹⁷¹ Zob. L. Bergström, *Imperatives and Ethics*, Stockholm University Press, Stockholm 1962.

¹⁷² Omówienie tego problemu znaleźć można w: H. N. Castañeda, *Imperative Reasonings...*, dz. cyt., s. 21-49.

¹⁷³ Jest to pogląd, którego szczegółowy opis znaleźć można w artykule A. Ros-
sa, *Imperatives...*, dz. cyt., s. 30-46. Zdaniem Castañedy stanowiska tego bronił P. T. Geach; P. T. Geach, *Imperative and Deontic Logic*, „Analysis” (18) 1958, s. 49-56. Krytykę tego poglądu znaleźć można w: H. N. Castañeda, *Imperatives and Deontic Logic*, „Analysis” (19) 1959, s. 42-48.

nione wnioskowanie byłoby poprawne. Zdaniem H. N. Castañedy, w logice spełniania generalne kryterium poprawności wnioskowań można przedstawić w następujący sposób:

Wnioskowanie jest poprawne, jeśli konkluzji automatycznie zostaje przypisana wartość prawdy czy spełnienia zawsze wtedy, gdy przesłanki posiadają wartość prawdy lub spełnienia, w zależności od tego, czy są one oznajmieniami, czy imperatywami¹⁷⁴.

Np. jeśli przesłanka „Zrób *a* i zrób *b*” jest spełniona, możemy stąd wywnioskować, że konkluzja „Zrób *a*” również jest spełniona.

Istnieją jednak problemy związane z powyższym podejściem. Krytykami takiej logiki byli np. A. F. Peters¹⁷⁵ i A. Ross¹⁷⁶. Według ostatniego z nich, nie możemy twierdzić, że logika spełniania i klasyczny rachunek zdań są paralelne, gdyż następujące wnioskowania, formalnie poprawne w świetle logiki klasycznej, jawią się jako paradoksalne¹⁷⁷:

V Wyślij list.
Wyślij list lub spal go.

Również Hare, zdając sobie sprawę z tego problemu, podał przykłady dyskusyjnych rozumowań, np.¹⁷⁸:

VI Włóż swój spadochron i wyskocz.
Wyskocz.

Wnioskowanie V wydaje się paradoksalne, gdyż osoba, która dostała rozkaz „Wyślij list”, może wywnioskować, iż spełni ów rozkaz poprzez jego spalenie. Również rozumowanie VI powoduje, że ktoś

¹⁷⁴ Tenże, *Imperative Reasonings...*, dz. cyt., s. 27. Opis tego stanowiska dokonany przez Hare’a, znaleźć można w: R. M. Hare, *Practical...*, dz. cyt., s. 62-63.

¹⁷⁵ Zob. A. F. Peters, *Hare on Imperative Sentences: A Criticism*, „Mind” (232) 1949, s. 535-540.

¹⁷⁶ A. Ross, *Imperatives...*, dz. cyt., s. 30-46. Zarzuty skierowane w stronę Hare’a logiki imperatywów znaleźć można również w: H. Zellner, *A Note on R. M. Hare and the Paradox of the Good Samaritan*, „Mind” (326) 1973, s. 281-282.

¹⁷⁷ Tamże, s. 38.

¹⁷⁸ Por. R. M. Hare, *Practical...*, dz. cyt., s. 65.

mógłby pomyśleć, że może wykonać rozkaz „Włóż swój spadochron i wyskocz” jedynie przez spełnienie nakazu „Wyskocz”.

W związku z powyższymi dylematami Hare obrał inną drogę wyjaśnienia praw logicznych. W *The Language of Morals* próbował oprzeć logikę imperatywów, a także innych wnioskowań, na pojęciach akceptacji, nieakceptacji (*assent i dissent*) oraz niezrozumienia (*misunderstanding*). Według niego,

Powiedzieć, że jeden sąd pociąga za sobą drugi, oznacza, iż nie możesz akceptować (*assent to*) pierwszego z nich i nieakceptować drugiego, chyba, że któregoś z nich nie zrozumiesz, przy czym ta „niemożność” jest „niemożnością” logiczną¹⁷⁹.

W późniejszym okresie tę samą definicję przedstawił w języku aktów mowy: „[...] akt mowy «p» pociąga akt mowy «q» wtedy i tylko wtedy, gdy «p» jest logicznie niespójne z negacją «q»: ktokolwiek akceptuje «p» ale neguje «q», zaprzecza sam sobie”¹⁸⁰. **Akceptację** (*assent to*) Hare zdefiniował następująco¹⁸¹: szczerze akceptujemy dane stwierdzenie wtedy i tylko wtedy, gdy wierzymy, że jest ono prawdziwe (jest tautologią, że nie możemy szczerze zaakceptować stwierdzenia, nie wierząc w nie w tym samym czasie¹⁸²); szczerze akceptujemy jakiś nakaz wtedy i tylko wtedy, gdy robimy lub postanawiamy zrobić to, co nadawca nam powiedział, że mamy zrobić (jest tautologią powiedzenie, że nie możemy szczerze zaakceptować nakazu, nie wykonując go w tym samym czasie¹⁸³). Kryterium poprawnego wnioskowania w logice akceptacji przedstawić więc można w następujący sposób:

Wnioskowanie z użyciem imperatywów jest poprawne, jeśli wszystkie przesłanki oraz konkluzja nie zostały niezrozumiane (zostały zrozumiane?), oraz wniosek jest akceptowany zawsze wtedy, gdy akceptowane są przesłanki¹⁸⁴.

Powyższa interpretacja może budzić pewne zastrzeżenia. Zarzucić jej można sprowadzanie logiki do psychologii¹⁸⁵. W takim wy-

¹⁷⁹ Tenże, *The Language...*, dz. cyt., s. 172.

¹⁸⁰ Tenże, *Foundationalism and Coherentism*, w: tenże, *Objective...*, s. 118.

¹⁸¹ Zob. tenże, *The Language...*, dz. cyt., s. 19-20.

¹⁸² Por. tamże, s. 20.

¹⁸³ Por. tamże.

¹⁸⁴ H. N. Castañeda, *Imperative Reasonings...*, dz. cyt., s. 33.

¹⁸⁵ Inne zastrzeżenie przedstawił H. N. Castañeda. Postawił on Hare’owi zarzut błędnego koła, ponieważ Hare zdefiniował, czym jest wynikanie, powołując się na logiczną konieczność. Zob. tamże, s. 29-30.

padku implikacja nie byłaby relacją analityczną, nie odnosiłaby się do logicznych wartości zdań, lecz do subiektywnych stanów przekonaniowo-wolicjonalnych podmiotu. Takie stany jednak nie podlegają prawu niesprzeczności i spójności. Prawa logiczne w tej interpretacji miałyby więc charakter przygodny, a nie konieczny. Wydaje się jednak, iż w teorii Hare'a zasady logiki opierają się na konwencjonalnych własnościach neustika, którego wyrazem jest akceptacja; neustik bowiem nie jest znakiem aktu psychicznego. „Akceptacja” czy „podpisywanie” się są słowami wyrażającymi zaangażowanie w daną czynność językową, zatem powyższe prawa można nazwać „logiką neustika”. Słowa „podpisać się” czy „akceptować” Hare nie używał do wyrażenia subiektywnego stanu umysłowego, lecz do wykonania konwencjonalnego aktu komunikacji. Podpisanie przeze mnie czeku – choć jest wyrazem zaangażowania – nie jest aktem mentalnym, bo aby podpis był ważny, taki akt nie musi mu towarzyszyć¹⁸⁶. Według Hare'a¹⁸⁷, prawa rządzące zaangażowaniem czy akceptacją powinniśmy traktować jako trzecią klasę konwencji, którą możemy dodać do przedstawionych przez Austina konwencji opisu i wskazywania¹⁸⁸. Konwencje subskrypcyjne¹⁸⁹ (*subscriptive conventions*) – bo o nich tu mowa – regulują sposób zaangażowania nadawcy w daną czynność językową. Odpowiadają one np. za to, że w sądzie, zeznając pod przysięgą, nie mógłbym powiedzieć, że widziałem, jak więzień zabija policjanta, a później twierdzić, iż jedynie żartowałem, lub wypisać, podpisać i wręczyć czek, a później stwierdzić, że udawałem, zatem jest on nieważny. To za sprawą konwencji subskrypcyjnych nie mogę również akceptować „ $p \sqcap q$ ” oraz „ p ”, nie akceptując zarazem „ q ”. Zatem Hare'a definicję wynikania należałoby zapisać przy użyciu znaku neustika (\vdash)¹⁹⁰:

¹⁸⁶ Por. R. M. Hare, *Some Subatomic...*, dz. cyt., s. 32.

¹⁸⁷ Zob. tamże, s. 37.

¹⁸⁸ „Konwencje opisu, korelujące słowa (= zdania) z typami sytuacji, rzeczy, zdarzeń itd., które można znaleźć w świecie. Konwencje wskazywania, korelujące słowa (= twierdzenia) z sytuacjami historycznymi, które można znaleźć w świecie. Mówi się o twierdzeniu, że jest prawdziwe, gdy historyczny stan rzeczy, z którym jest ono skorelowane za pomocą konwencji wskazywania (stan rzeczy, do którego «się odnosi»), jest takim typem stanu rzeczy, z którym zdanie użyte do wypowiedzenia go jest skorelowane za pomocą konwencji opisu”; J. L. Austin, *Mówienie...*, dz. cyt., s. 162-163.

¹⁸⁹ Zob. R. M. Hare, *Some Subatomic...*, dz. cyt., s. 37.

¹⁹⁰ Wszystkie prawa formalne moralnej gry językowej Hare'a należy przedsta-

Definicja wynikania

$$\vdash (p \sqcap q) = \text{df } \sim (\vdash p ? \sim \vdash q)$$

Zdaniem Castañedy, powyższe rozwiązanie tylko z pozoru unieważnia problemy logiki spełniania, ponieważ po podstawieniu za pojęcie „akceptacji” tego, przez co Hare je definiuje, okaże się, że to podejście zredukuje się do logiki spełniania¹⁹¹. Logika spełniania natomiast podlega wielu trudnościom wzmiankowanym powyżej. W późniejszych latach swego życia Hare jej jednak bronił¹⁹². Jego zdaniem, problemy wskazane przez A. Rossa dają się wyjaśnić i ominąć. Inferencja z „Wyślij list” do „Wyślij lub spal list” wydaje się paradoksalna, ponieważ implikatury konwersacyjne¹⁹³ zdania pierwszego i drugiego bardzo się od siebie różnią¹⁹⁴. Gdy zwracamy uwagę na treść rozumowania, podobne paradoksy możemy zaobserwować również wśród stwierdzeń: z „Sokrates jest śmiertelny” możemy wywnioskować „Sokrates jest śmiertelny lub leży nad Adriatykiem”. Co więcej, mylimy się, jeśli uważamy, iż logika spełniania zakłada, że spełnienie wniosku rozumowania jest z konieczności spełnieniem jego prze-

wiać przy użyciu znaku „ \vdash ”; z tego też powodu zasadę wynikania imperatywów jednostkowych czy sprzeczności imperatywów prezentowaliśmy, stosując symbol „ \vdash ”.

¹⁹¹ Zob. H. N. Castañeda, *Imperative Reasonings...*, dz. cyt., s. 33.

¹⁹² Zob. R. M. Hare, *Some Alleged...*, dz. cyt., s. 42.

¹⁹³ „Termin ukuty przez Grice'a na oznaczenie implikacji wypowiedzi wykraczających poza wnioski w ścisłym sensie wynikające z treści. Gdy na pytanie, co sądzę o moim nowym koledze, odpowiadam: «Powiadają, że ma niezłą wymowę», to z tego, co powiedziałem, nie wynika wprawdzie logicznie nic na temat jego zdolności filozoficznych, ale moje słowa implikują, że mam o tych zdolnościach nie najlepsze zdanie. Implikatury mogą być zasadniczym nerwem rozmowy, np. gdy potępia się przez zdawkowe pochwały. Ocena znaczenia słów zależy często od tego, które implikacje wypowiedzi zawierających te słowa uznamy za implikatury, a które za rzetelne konsekwencje logiczne tego, co się mówi. Implikatury można podzielić na dwa rodzaje: zilustrowane wyżej implikatury konwersacyjne i bardziej subtelne implikatury konwencjonalne. W tym wypadku jedno wyrażenie może mieć taką samą teść jak inne, a związana z jego użyciem implikatura może być sprawą konwencji. Niektórzy różnicę między «on jest biedny i uczciwy» a «on jest biedny, ale uczciwy» objaśniają w ten sposób, że choć oba wyrażenia mają taką samą treść (są prawdziwe w ściśle tych samych warunkach), to za sprawą znaczenia słówka «ale» drugie ma implikatury (owo połączenie jest zaskakujące czy godne uwagi) jakich brak pierwszemu”; S. Blackburn, *Implikatura*, w: S. Blackburn, *Oksfordzki...*, dz. cyt., s. 168.

¹⁹⁴ Zob. R. M. Hare, *Some Alleged...*, dz. cyt., s. 29.

słanek¹⁹⁵. Jeśli C jest nakazem, a $c_1, c_2 \dots c_n$ nakazami, które mogą być wydedukowane z C, nie możemy wykonać C, jeśli jedynie wykonamy c_1 ¹⁹⁶. Prawa logiki imperatywów rozpadają się bowiem na *rozumowania do warunku koniecznego* i *rozumowania do warunku wystarczającego*. Jeśli powiedziano nam, abyśmy włożyli spadochron i wyskoczyli, a my wyciągniemy stąd wniosek, że musimy wyskoczyć, nie będzie to wniosek błędny. Dedukcja taka będzie poprawna. Błędne będzie natomiast nieuzasadnione przekonanie, że jest to wszystko, co musimy zrobić. Zakładamy bowiem, że wniosek („Wyskocz!”) dostarcza wystarczających warunków, aby spełnić przesłankę („Włóż spadochron i wyskocz!”), podczas gdy dostarcza on jedynie warunku koniecznego¹⁹⁷. Prawa logiki warunków koniecznych i wystarczających są izomorficzne z prawami klasycznego rachunku zdań¹⁹⁸, gdyż jeśli p pociąga za sobą q (w zwykłym sensie tego słowa), to q jest warunkiem koniecznym prawdziwości czy spełnienia p , p natomiast jest warunkiem wystarczającym prawdziwości czy spełnienia q . Jeśli mamy powiązane ze sobą dwa imperatywy i_1 i i_2 w taki sposób, że zdanie „ i_1 jest spełniony” pociąga za sobą „ i_2 jest spełniony”, to osoba, która poda i_1 i zastanawia się, jak go spełnić, może wywnioskować, że spełnienie i_2 będzie koniecznym warunkiem spełnienia i_1 . Natomiast osoba, która poda i_2 i zastanawia się, jak go spełnić, może wywnioskować, iż spełnienie i_1 będzie warunkiem wystarczającym spełnienia i_2 ¹⁹⁹.

Zarówno logikę imperatywów, jak i izomorficzną z nią logikę oznajmień Hare próbował opisać za pomocą terminów niezależnych od pojęcia prawdy. Jak twierdził, logika jest zbiorem zasad opisujących i kierujących poprawnymi sposobami wnioskowania bez względu na to, czy przesłanki są prawdziwe, fałszywe, czy nie posiadają owych wartości²⁰⁰. Nie interesował się więc wyjaśnieniem semantyki imperatywów. Pisał:

Zarzykuję stwierdzenie, że ponieważ nie istnieje nic, co mogłoby odpowiadać prawdzie, gdy mówimy o imperatywach, ci, którzy poszukują takiej na-

¹⁹⁵ Zob. tamże, s. 30.

¹⁹⁶ Por. tenże, *Imperative Sentences...*, dz. cyt., s. 32.

¹⁹⁷ Por. tenże, *Some Alleged...*, dz. cyt., s. 30.

¹⁹⁸ Zob. tenże, *Practical...*, dz. cyt., s. 67.

¹⁹⁹ Zob. tamże.

²⁰⁰ Zob. tenże, *Imperative Sentences...*, dz. cyt., s. 18.

miastki prawdy, aby ustalić podstawy logiki imperatywów, szukają w złym miejscu²⁰¹.

§ 4.12 Problem przejścia od „jest” do „powinien”

Bardzo ważny wkład w rozwój metaetyki wniosło podjęcie na początku XX wieku wątpliwości podniesionej przez Davida Hume’a, tzw. problemu przejścia od „jest” do „powinien”. W *Traktacie o naturze ludzkiej* wysunął on pytanie, na które wielu etyków ze szkoły analitycznej czuło się w obowiązku odpowiedzieć. Również Richard Hare zmierzył się z tym dylematem. Ten, jak twierdzi T. Styczeń²⁰², najważniejszy dla moralnej filozofii anglosaskiej fragment tekstu, brzmi następująco:

W każdym systemie moralności, z jakim dotychczas się spotykałem, stwierdzałem zawsze, że autor przez pewien czas idzie zwykłą drogą rozumowania, ustala istnienie Boga, albo robi spostrzeżenia dotyczące spraw ludzkich; aż nagle nieoczekiwanie i ze zdziwieniem znajduję, iż zamiast zwykłych spójek, jakie znajduje się w zdaniach, a mianowicie jest i nie jest, nie spotykam żadnego zdania, które by nie było powiązane słowem *powinien* albo *nie powinien*. Ta zmiana jest niedostrzegalna, lecz niemniej ma wielką doniosłość. Wobec tego bowiem, że to *powinien* albo *nie powinien* jest wyrazem pewnego nowego stosunku czy twierdzenia, przeto jest rzeczą konieczną te zwroty zauważyć i wyjaśnić, a jednocześnie konieczne jest, iżby została wskazana racja tego, co wydaje się całkiem niezrozumiałe, a mianowicie jak ten nowy stosunek może być wydedukowany z innych stosunków, które są całkiem różne od niego²⁰³.

Wywołany przez Hume’a spór o relację pomiędzy powinnością a opisem oraz ewentualny jej charakter prowadzi do doniosłych konsekwencji praktycznych. Przyjmowane rozstrzygnięcia mają duży wpływ na stanowisko w sprawie metody uzasadniania sądów etycznych. Od nich zależy w dużej mierze to, czy etykę uznać powinniśmy za dziedzinę autonomiczną wobec nauk przyrodniczych, tzn. czy metoda uzasadniania jej zasad jest podobna do metody obowiązującej w naukach, czy raczej od niej różna.

Istnieje wiele interpretacji postawionej przez Hume’a wątpliwości²⁰⁴. Generalnie rzecz ujmując, fragment ten u różnych komenta-

²⁰¹ Por. tenże, *Some Alleged...*, dz. cyt., s. 42-43.

²⁰² T. Styczeń, *Spór o naukowość etyki*, w: I. Lazari-Pawłowska (red.), *Metaetyka...*, dz. cyt., s. 557.

²⁰³ D. Hume, *Traktat...*, dz. cyt., s. 259-260.

²⁰⁴ Zob. M. Rutkowski, *Rola...*, dz. cyt., s. 195-216.

torów przybierał postać bądź ontologicznego twierdzenia o braku jakiegokolwiek związku między obowiązkiem a faktami, bądź semantycznej formuły o niemożliwości logicznego przejścia od sądów opisowych do powinnościowych. Ponieważ Hare stronił od ontologii, pierwszą interpretacją nie będziemy się zajmowali. Ważne jest natomiast to, że można wyróżnić kilka podstawowych rodzajów interpretacji semantycznej²⁰⁵. Według najbardziej popularnej, musimy przystać na fakt, iż pomiędzy zdaniem orzekającym a powinnościowymi nie istnieją żadne relacje, które umożliwiłyby poprawne wyprowadzenie zdań powinnościowych z orzekających. W konsekwencji uniemożliwia to uzasadnienie sądów moralnych na podstawie jakiegokolwiek zbioru sądów niemających moralnego charakteru. Przedstawiciele tego poglądu nazwać można zwolennikami **autonomii całkowitej**²⁰⁶; głoszą oni, że pomiędzy zdaniem orzekającym a powinnościowymi nie istnieją żadne relacje umożliwiające poprawne rozumowanie. Drugą grupę stanowią zwolennicy **autonomii formalnologicznej**²⁰⁷. Według nich, inferencja od zdań orzekających do powinnościowych nie jest niemożliwa, lecz aby się dokonała, potrzebne jest założenie we wnioskowaniu dodatkowych przesłanek, które by ją uprawomocniły. Sformułowana w tym zakresie autonomia oznacza, że zdania powinnościowe nie mogą stanowić wniosku rozumowania dedukcyjnego, zawierającego w roli przesłanek jedynie zdania orzekające. Wnioskowanie takie może być poprawne, jedynie gdy: (a) w przesłankach występują terminy zawarte w konkluzji; lub (b) są one wyrazem uczuć wypowiedzianego je podmiotu lub (c) wśród przesłanek są również oceny moralne²⁰⁸.

Sposób odpowiadania na *is/ought problem* często uzależniony jest od tradycji reprezentowanej przez danego filozofa. Hare, jako spadkobierca zapoczątkowanego przez Hume'a nurtu emotywistycznego, utrzymywał, iż uwaga zawarta w *Traktacie* ma druzgocące konsekwencje dla wszelkich deskrytywistycznych stanowisk etycznych, przy czym traktował ją jako tezę potwierdzającą autonomię formalnologiczną. Interesujące nas przejście będzie niemożliwe, jeśli wśród przesłanek nie będzie zalecenia. Hare odrzucał stanowi-

sko, które można nazwać mocną logiczną interpretacją tezy Hume'a. Sądy deskrytywne mają i faktycznie są ważnym elementem dedukcji moralnych. Stanowią one również racje przyjęcia sądów oceniających. Jednak ilekroć przechodzimy od faktu do powinności („Powinienniśmy pomóc temu człowiekowi, ponieważ jest głodny”), jest w tym rozumowaniu luka ukrywająca przesłankę większą („Powinienniśmy pomagać głodującym”). Ta większa przesłanka sama może figurować jako wniosek innego sylogizmu, ale w pewnym momencie łańcuch rozumowań musi się kończyć na zasadzie, której nie potrafimy uzasadnić za pomocą faktów oraz innej zasady moralnej. Ma ona bowiem charakter niesprowadzalnego do opisu, niezdeternowanego surowymi faktami zalecenia²⁰⁹.

Zdaniem Hare'a, następujące dwie zasady są niepodważalne:

(A) „Żaden wniosek imperatywny nie może być poprawnie wyprowadzony ze zbioru przesłanek, które nie zawierają przynajmniej jednego imperatywu”²¹⁰.

(B) „Żaden wniosek oznajmujący nie może być poprawnie wyprowadzony ze zbioru przesłanek, które nie mogłyby być poprawnie wyprowadzone jedynie ze zdań oznajmujących”²¹¹.

Hare uzasadniał tezę (A), odwołując się do preskrytywnej funkcji języka etycznego i zmodyfikowanej wersji *testu otwartego pytania* Moore'a. Była ona jednak dla niego na tyle oczywista, iż twierdził, że zasada (A) jest chyba najprostszym i najważniejszym stwierdzeniem w etyce²¹². W okresie, w którym pisał *The Language of Morals*, wśród filozofów panowała bowiem raczej powszechna zgoda co do trafności omawianych tez. Dlatego też Hare nie przedstawił systematycznego wykładu uzasadniającego akceptację powyższych twierdzeń. Uwierzytelniał je jedynie, podając przyczyny słabości naturalizmu, nie zaproponował jednak pozytywnych argumentów na ich poparcie. W okresie późniejszym, w którym tacy filozofowie, jak Phillipa Foot, Peter Geach, John Searle czy Alasdair MacIntyre, próbowali odrodzić, jak się zdawało Hare'owi, martwy deskrytywizm, pisał on natomiast polemiczne eseje przeciwko ich poglądom. Powyższe dwie zasady traktował więc jako dogmat logiki, którą rozwijali tacy przed-

²⁰⁵ Por. tamże, s. 203.

²⁰⁶ Zob. tamże, s. 197-199.

²⁰⁷ Por. tamże, s. 203-204.

²⁰⁸ Zob. tamże, s. 203.

²⁰⁹ Por. R. M. Hare, *The Language...*, dz. cyt., s. 46.

²¹⁰ Tamże, s. 28.

²¹¹ Tamże.

²¹² Zob. tamże, s. 31.

stawiciele filozofii analitycznej, jak B. Russell czy R. Carnap²¹³. Twierdził, iż każde uzasadnienie czy wnioskanie, którego konkluzją jest sąd o charakterze moralnym, musi się opierać na dedukcji, a każda taka inferencja dedukcyjna musi mieć charakter analityczny²¹⁴: nie możemy wydedukować z przesłanek niczego, co nie byłoby w nich zawarte choćby *implicite*. Zasadę tę nazwijmy **zasadą analityczności dedukcji** (we wniosku nie może być zawarte nic więcej ponad to, co jest w przesłankach, z wyjątkiem tego, co może być dodane na mocy definicji znaczeń wyrażań w nich występujących). Zasadę tę interpretował w praktyce jako mówiącą, że aby wyprowadzić tryb imperatywny, przynajmniej jedna z przesłanek musi być preskrypcją²¹⁵. Hare odcinał się więc od rozwiązań, które proponował S. E. Toulmin²¹⁶, tzw. inferencji ewaluatywnej, która rządziłaby się innymi zasadami, mniej restrykcyjnymi niż dedukcja²¹⁷.

Powyższe uzasadnienie (zasadę analityczności dedukcji) poddać można jednak w wątpliwość, istnieją bowiem poprawne wnioskania, które polegają na wyprowadzeniu sądu opisowego z imperatywu, np.:

Nakarm tego kota!

Uważam, że to jest kot.

Jeśli powyższy przykład skonceptualizujemy jako przejście z tropika imperatywnego do stwierdzającego na mocy zawartego w obu zdaniach frastika, to uzasadnienie tez (A) i (B) nie może się opierać

²¹³ Zarzut bezkrytycznego posługiwania się przez Hare'a dystynkcją analityczne – syntetyczne, która była wykorzystywana przez niego w argumentowaniu przeciwko deskryptywizmowi, znaleźć można w: B. A. McKinnon, *Hare's Use of Hume's Fork*, „Ethics” (84) 174, s. 332-338.

²¹⁴ Zob. tamże, s. 32, 33.

²¹⁵ Zob. tamże, s. 32.

²¹⁶ Zob. S. E. Toulmin, *An Examination of the Place of Reason in Ethics*, Cambridge University Press, Cambridge 1958, s. 38.

²¹⁷ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 45-53. Hare podaje trzy argumenty przeciwko koncepcji słabej inferencji (*loose inference*) Toulmina:

a) specjalne zasady wnioskania, słaba inferencja nie dotyczy formy słów i zdań, ale ich treści: reguły owe są bardziej podobne do reguł naukowych niż logicznych; b) sprowadzanie sporów empirycznych do sporów lingwistycznych przeczy natomiast naszym intuicjom. Nie jest jasne, czym jest owa słaba inferencja, na czym ma ona polegać; c) największym jednak błędem takiego podejścia jest pozbawienie moralnego myślenia kluczowego elementu: decyzji.

na zasadzie analityczności dedukcji, skoro wnioskanie od imperatywu do stwierdzenia jest poprawne. W przesłance bowiem nie ma tropika odpowiadającego za stwierdzenie, pojawia się on natomiast we wniosku.

W podobny sposób tezę (A) skrytykował John Searle, który zinterpretował ją w duchu teorii czynności mowy i nazwał **błędem błędu naturalistycznego**²¹⁸. Jeśli strukturę głęboką sądów moralnych opiszemy nie poprzez ich analogię do imperatywów, ale specyficzne czynności mowy, które z konieczności je charakteryzują, tezę (A) powinniśmy przedstawić jako przekonanie, iż żadne zalecenie nie może być poprawnie wydedukowane z przesłanek, jeśli chociaż jedna z nich nie ma preskryptywnej mocy illokucyjnej. Według Searle'a,

u podłoża błędu naturalistycznego błędu leży fundamentalna zasada teorii języka, w myśl której istnieje przepaść logiczna między znaczeniem wyrażenia wartościującego a kryteriami jego stosowania²¹⁹.

*Błąd czynności mowy*²²⁰ pociąga więc *błąd błędu naturalistycznego*²²¹. Jeśli bowiem znaczenie nie ma związku z czynnością illokucyjną, którą wykonujemy wypowiadając dane zdanie, oraz stwierdzimy, że wyznaczenie jest sprawą znaczenia,

twierdzenie P, wygłoszone, gdy wypowiadamy zdanie S, może pociągać twierdzenie Q, wygłoszone, gdy wypowiadamy zdanie T, nawet gdy wypowiadanie S ma w sposób swoisty pewną moc illokucyjną, a wypowiadanie T ma inną moc²²².

Istotą *błędu błędu naturalistycznego* jest więc

wyprowadzenie z faktu, że dwie wypowiedzi mają różne moce illokucyjne, tego, że zdanie wyrażone w jednej wypowiedzi nie może pociągać zdania wyrażonego w drugiej²²³.

To, iż zdania mają różne moce illokucyjne, nie stoi na przeszkodzie, aby wchodziły one z sobą w relacje inferencyjne ze względu na treść, którą niosą.

Zauważyć można, że krytyka ta zakłada, iż filozofowie, którzy uważają, że logika niedeskryptywnych czynności mowy (np. logika zale-

²¹⁸ Zob. J. Searle, *Czynności...*, dz. cyt., 168-174.

²¹⁹ Tamże, s. 171.

²²⁰ Zob. § 3.8.

²²¹ Zob. tamże, s. 178.

²²² Tamże, s. 195.

²²³ Tamże, s. 173.

ceń czy pytań) jest niemożliwa, są w błędzie. Zdaniem Searle'a, możliwe są bowiem inferencje pomiędzy wypowiedziami, które w myśl klasycznej teorii logicznej nie są sądami logicznymi. Jak pisze,

z faktu, że cel czy illokucyjna wypowiedziana zdania [gramatycznego] jest wartościująca, nie wynika, iż zdanie [logiczne] wyrażone nie może być następstwem zdania [logicznego] wyrażonego, gdy wypowiadamy zdanie [gramatyczne], którego moc illokucyjna czy cel wypowiedziane jest „opisowy”. Warunki prawdziwości jednego zdania [logicznego] mogą być wystarczające dla spełnienia warunków prawdziwości drugiego – nawet gdy cel wypowiedziania jednego zdania [gramatycznego] może różnić się od celu wypowiedziania drugiego²²⁴.

Richarda Hare'a i Johna Searle'a różni jednak odmienne rozumienie owej logiki aktów mowy. Obydwaj zaakceptowaliby bowiem rozumowanie o formie:

$$(1) \quad \frac{p \sqcup !q}{p} \\ !q$$

jednak jedynie Searle przyjąłby

$$(2) \quad \frac{p \sqcup q}{p} \\ !q$$

W kwestii tej pomiędzy Searle'em a Hare'em toczyła się burzliwa dyskusja. Spowodowana została artykułem pierwszego z nich *Jak wywieść „powinien” z „jest”*²²⁵, na który Hare odpowiedział tekstem *Gra w obiecywanie*²²⁶. Według Searle'a, osoba wypowiadająca zdanie typu: „Niniejszym obiecuję ci, Smith, zapłacić 5 dolarów”, na mocy znaczenia wypowiedzianych przez nią słów staje się zobowiązana i powinna uczynić to, co wypowiedziała. Stosując przekształcenia równoważnościowe, nie dodając żadnych przesłanek wartościujących, można bowiem wykazać, iż ze zdania: „Jones wypowiedział słowa: «Niniejszym obiecuję ci, Smith, zapłacić 5 dola-

²²⁴ Tamże, s. 187-188.

²²⁵ Tenże, *Jak wywieść „powinien” z „jest”*, tłum. J. Hołówka, „Etyka” (16) 1978, s. 163-177.

²²⁶ R. M. Hare, *Gra...*, dz. cyt., s. 151-164.

rów» wynika: „Jones powinien zapłacić 5 dolarów Smithowi”²²⁷. Aby to udowodnić, Searle odwołał się do reguł konstytutywnych²²⁸, które głoszą, że kto składa obietnicę, a składa ją ten, kto wykonuje performatyw: „Niniejszym obiecuję, że...”, bierze na siebie zobowiązanie, decydując bowiem się na pewną grę językową, musi przestrzegać jej reguł. Reguły te mówią natomiast, że każdy, kto się zobowiąże, powinien dotrzymać zobowiązania. Zdaniem Searle'a, owe zespoły reguł, składają się m.in. z faktów instytucjonalnych, tworzą tzw. instytucję obiecywania²²⁹. Ponieważ w powyższym rozumowaniu nie odwołujemy się do zdań wartościujących, a jedynie do pewnych usankcjonowanych instytucji językowych, reguł konstytutywnych, które tworzą instytucjonalne fakty, możemy wnioskować, że istnieją dziedziny, w których możemy „przejsć od jest do powinien”, gdyż pewne systemy reguł konstytutywnych wiążą się z obowiązkami, zobowiązaniami i odpowiedzialnością.

Zdaniem Hare'a, zaprezentowane przez Searle'a wnioskowanie nie jest jednak dowodem na możliwość takiego przejścia, przemyczone są w nim bowiem ukryte przesłanki wartościujące. Nie jest tak, jak chciałby Searle, że wchodzące w skład instytucji obiecywania reguły konstytutywne są koniecznymi prawdami językowymi. Według Hare'a,

jeżeli mamy przyjąć owe wyrażenia performatywne, to musimy również przyjąć pewne konstytutywne reguły syntetyczne (a nie tylko językowe), tworząc w ten sposób określoną instytucję, w której wyrażenia te mają sens²³⁰.

Reguły konstytutywne są regułami syntetycznymi, fakty instytucjonalne zatem zawierają w sobie element wcześniejszego ich ustanowienia na podstawie woli, rozumu lub też arbitralnych decyzji. Mają więc swe źródło w normatywnym wyborze. Aby dana praktyka językowa (np. wypowiedzenie „obiecuję, że...”) pociągała za sobą

²²⁷ Por. J. Searle, *Jak wywieść...*, dz. cyt., s. 164.

²²⁸ „Reguły regulatywne regulują formy zachowań istniejących uprzednio lub niezależnie, [...] reguły konstytutywne nie tylko regulują; one tworzą lub określają nowe formy zachowania. [...] Reguły konstytutywne konstytuują (a również regulują) działalność, której istnienie jest logicznie zależne od reguł”; tenże, *Czynności...*, dz. cyt., s. 49.

²²⁹ „«Instytucje» te są systemami reguł konstytutywnych. U podłoża każdego faktu instytucjonalnego leży reguła (system reguł) o postaci: «X uchodzi za Y w kontekście C»; tamże, s. 71.

²³⁰ R. M. Hare, *Gra...*, dz. cyt., s. 151-164.

faktyczne zobowiązania, wcześniej musimy przyjąć regułę syntezy, która mówi, że właśnie tak ma się dziać. To, jakie reguły powinniśmy przyjąć, podlega jednak decyzji moralnej. Z tej perspektywy Searle, odwołując się do pewnych instytucji, odnosi się do zastanego zbioru obowiązków *prima facie*. Zadaniem etyki jest jednak przede wszystkim podejmowanie racjonalnych wyborów dotyczących charakteru instytucji oraz reguł konstytutywnych, które powinny obowiązywać, a tego typu decyzje nie mogą być zdeterminowane wyłącznie surowymi faktami. Faktów instytucjonalnych nie można zredukować do faktów surowych²³¹, pierwsze z nich bowiem to zakamuflowane sądy normatywne. Dla Searle'a to,

czy cała instytucja obiecywania jest dobra, czy zła, czy zobowiązania podjęte, gdy obiecyjemy, zostają usunięte przez inne zewnętrzne względy – to pytania zewnętrzne wobec samej instytucji²³².

Dla Hare'a natomiast są to najważniejsze pytania etyczne i musimy dysponować odpowiednią teorią i językiem, który pomogą nam je rozstrzygnąć. Argument Searle'a nie dowodzi, że możemy ocenić instytucję obiecywania, normatywną zasadę *prima facie* typu B, odwołując się do reguły wnioskowania (2). Jeśli chcemy przyjąć, że inferencja

$$(2) \quad \frac{p \sqsupset q}{p} \quad \frac{p}{!q}$$

jest prawidłowa, wcześniej musimy założyć istnienie odpowiednich syntetycznych reguł konstytutywnych, które wiążą ze sobą odmienne moce illokucyjne i stanowią ukryte przesłanki wartościujące. Ponieważ reguł konstytutywnych nie możemy wyprowadzać przy użyciu powyższego prawa logicznego, procedura taka prowadzi bowiem do błędnego koła, najważniejsze wnioski moralne – uzasadniające przyjmowane zasady etyczne – nie polegają na dedukcji sądów powinnościowych z przesłanek opisowych.

²³¹ Zdaje sobie z tego sprawę również Searle: „Nie usiłuję sprowadzać faktów instytucjonalnych do faktów surowych, nie ma więc w analizie motywacji redukcjonistycznej”. Zob. J. Searle, *Czynności...*, dz. cyt., s. 76.

²³² Tamże, s. 234.

§ 4.13 Internalizm R. M. Hare'a

Kolejną istotną cechą wszystkich zaleceń, która współtworzy to, co Hare nazwał logiką języka moralnego, jest ich specyficzny związek z motywacją do działania. W dziedzinie moralności jest on powodem akceptacji przez Hare'a internalizmu motywacyjnego. Najogólniej rzecz ujmując, przedstawiciele **internalizmu motywacyjnego** uważają, że motyw do działania są „wewnętrzne” wobec obowiązku czy moralnej powinności. Innymi słowy, jest to pogląd, że istnieje jakiś typ koniecznej relacji występującej pomiędzy sądami moralnymi a motywem do działania. **Eksternaliści** temu zaprzeczają²³³. Istnieje wiele odmian internalizmu. Dzieleny jest on przede wszystkim ze względu na niejednolity sposób pojmowania motywu, obowiązku oraz relacji pomiędzy motywem a obowiązkiem. Różnym jego wersjom nadano odmienne nazwy²³⁴. W celu opisanego internalizmu Hare'a posłużymy się podziałem na trzy płaszczyzny, w ramach których rozgrywa się spór pomiędzy internalizmem motywacyjnym a eksternalizmem. Wyróżnić można:

(I) **Płaszczyznę ontologiczną** – dotyczy ona własności, które mogą wyznaczać powinność. Z tej perspektywy można mówić o realizmie czy antyrealizmie racji do działania (realizmie i antyrealizmie motywacyjnym²³⁵) i rozstrzygać, jakiego rodzaju własności im przysługują.

(II) **Płaszczyznę semantyczną** – pozwala ona na uchwycenie związków znaczeniowych pomiędzy moralną powinnością (normą etyczną), a motywem czy racją do działania. Na tym poziomie możemy mówić np. o analityczności lub syntetyczności, konieczności lub przygodności zdań opisujących relacje występujące pomiędzy znaczeniami powyższych terminów.

(III) **Płaszczyznę psychologiczną** – tutaj ustala się naturę samego motywu, np. funkcję motywacyjną przyznać można bądź pojmowa-

²³³ Do współczesnych eksternalistów motywacyjnych zalicza się: Williama Frankę, Philippę Foot, Thomasa Scanlona, Petera Railtona, Davida Brinka.

²³⁴ M. Rutkowski opisuje następujące rodzaje internalizmu: internalizm racji, pragnień, mocny, słaby, subiektywny, obiektywny, transcendentalny, epistemiczny, postawy, klasyczny, typu kognitywistycznego, typu nonkognitywistycznego. Zob. M. Rutkowski, *Obowiązek...*, dz. cyt.

²³⁵ *Realizmem motywacyjnym* nazwiemy pogląd, głoszący, iż istnieją realne własności moralne, które motywują nas do działania. Teorię taką zakłada np. analizowany przez M. Rutkowskiego *internalizm transcendentalny*. Zob. tamże, s. 257. *Antyrealizm motywacyjny* temu zaprzecza.

nym w różny sposób pragnieniom bądź niezależnym od nich racjom do działania.

Ponieważ Richard Hare swój internalizm uzasadniał przede wszystkim, dowodząc, iż jego źródło tkwi w logice języka moralnego, skupimy się na płaszczyźnie semantycznej. Rozważania z poziomu psychologicznego posłużą natomiast do odparcia niektórych zarzutów stawianych Hare'owi oraz uzasadnienia przedstawionej poniżej interpretacji jego stanowiska. Najmniej uwagi poświęcimy poziomowi ontologicznemu, ponieważ Hare nie przykładał wielkiej wagi do analiz ontologicznych, uważając, że problemy związane z istnieniem podejmować należy na płaszczyźnie językowej. Wystarczy stwierdzić, że był on antyrealistą²³⁶. Odrzucił realizm motywacyjny, wierzył bowiem, iż skutecznie odpiera go argument Mackiego – tzw. *argument z dziwaczności*²³⁷.

Hare był przekonany, że o mocy motywacyjnej sądów moralnych decyduje fakt, iż są one zaleceniami. Preskrypcje – zatem również imperatywy i sądy moralne – motywują, ponieważ są wyrazem szeroko rozumianej woli osób je akceptujących. Szczera ich akceptacja staje się motywem do działania. Szczera akceptacja polega na przyjęciu neustika, podpisaniu się pod zdaniem, którego jesteśmy odbiorcą. Jedno z podstawowych praw logiki zaleceń – prawo internalizmu moralnego (PIM) – mówi:

Jeśli ktoś jest odbiorcą danego zalecenia oraz szczerze je zaakceptuje, musi mieć dyspozycję, aby zgodnie z nim postąpić, w przeciwnym wypadku oskarżyć go można o logiczną sprzeczność.²³⁸ „Wypowiedź preskryptywna jest taką wypowiedzią, której każdy na mocy pojęciowej konieczności musi podporządkować swoje działanie, jeśli szczerze ją przyjąć”²³⁹.

Aby podkreślić pojęciowy charakter swojego internalizmu, Hare przedstawił go w *The Language of Morals* poprzez opis koniecznych relacji występujących między sądami moralnymi a imperatywami. W związku z tym oprócz PIM1 wyróżnić można jeszcze cztery prawa pojęciowe, determinujące jego internalizm:

²³⁶ Zob. § 3.3.3.

²³⁷ Zob. § 3.3.2.

²³⁸ Zob. R. M. Hare, *Imperatives, Prescriptions...*, dz. cyt., s. 43.

²³⁹ Zob. tenże, *Internalism and Externalism in Ethics*, w: tenże, *Objective...*, dz. cyt., s. 101.

PIM2 Szczera akceptacja imperatywu jednostkowego „Wykonam x!” pociąga za sobą pojęciową konieczność posiadania motywu do wykonania x (wykonanie lub postanowienie wykonania nakazanego działania x)²⁴⁰.

PIM3 Wszystkie zalecenia pociągają za sobą imperatywy jednostkowe – jeśli mają być motywem do działania, muszą pociągać za sobą imperatywy²⁴¹. Jeśli szczerze akceptuję sąd „W sytuacji x powinno się pomóc słabszemu”, na mocy logicznej konieczności muszę zaakceptować imperatyw: „Pomogę słabszemu w sytuacji x!”, w przeciwnym razie oskarżyć mnie można o logiczną sprzeczność²⁴².

PIM3 jest bardzo ważną cechą języka preskryptywnego, odpowiada bowiem za to, że ten język nazywamy **normatywnym (ewaluatywnym)** a sądy moralne ocenami. Hare, odwołując się do tego prawa, w *The Language of Morals* podał następujące kryterium ewaluatywności sądów (KES):

KES Testem na to, czy ktoś używa sądu „Powiniem zrobić X” jako sądu wartościującego jest odpowiedź na pytanie «Czy rozpoznaje on, czy też nie, że jeśli zaakceptuje (*assent to*) powyższy sąd, musi również zaakceptować komendę „Niech zrobię X”²⁴³.

Aby sąd był ewaluatywny, osoba, która go akceptuje, musi zaakceptować imperatyw, który jest z niego wyprowadzalny. Według Hare'a, jest to prawdą na mocy definicji słowa ewaluatywne²⁴⁴, sądy opisowe same z siebie nie mogą bowiem motywować²⁴⁵.

Ponieważ sądy moralne są uniwersalnymi zaleceniami, z PIM2 i PIM3 wynika:

²⁴⁰ Zob. tenże, *The Language...*, dz. cyt., s. 20.

²⁴¹ Zob. tamże, s. 163. Krytykę tej tezy znaleźć można w: R. Hancock, *A Note on Hare's The Language of Morals*, „The Philosophical Quarterly” (50) 1963, s. 56-63.

²⁴² Odwołując się do wprowadzonych symboli oznaczających różne rodzaje zaleceń, prawo to można uogólnić i przedstawić w następujący sposób: jeśli osoba *O* podpisze się pod zdaniem *p*, które jest preskrypcją *A*, *A'*, *ZTI*, *B*, *C* lub *D*, na mocy konieczności pojęciowej musi ona zaakceptować zdanie *p*, które ma formę *A*.

²⁴³ R. M. Hare, *The Language...*, dz. cyt., s. 168.

²⁴⁴ Zob. tamże, s. 172.

²⁴⁵ Zob. tamże, s. 29, 44, 46.

PIM4 Szczera akceptacja sądu moralnego „Powinno się wykonać x” pociąga za sobą konieczność posiadania motywu aby wykonać x. Zaakceptowanie przez osobę O sądu moralnego, analitycznie pociąga za sobą konieczność przyjęcia zdania „Osoba O ma motyw”²⁴⁶.

Wielu komentatorów Hare'a, opierając się przede wszystkim na pracy *The Language of Morals*, sformułowało przekonania odbiegające od praw PIM1–PIM4. Znaleźć tam można bowiem fragmenty, w których Hare twierdzi, iż „jest tautologią, że nie możemy szczerze zaakceptować nakazu skierowanego w naszą stronę, i w *tym samym czasie* go nie wykonać [podkreślenie K. S.], jeśli właśnie mamy ku temu okazję oraz jest to w naszej (fizycznej i psychicznej) mocy”²⁴⁷. Sformułowanie to doprowadziło do przypisania mu **mocnej wersji internalizmu** (MI), uważa się bowiem, że według niego posiadanie moralnego motywu jest wystarczającym warunkiem samego działania – mając rację do działania, nie możemy działać w sposób z nią sprzeczny²⁴⁸.

MI Szczera akceptacja sądu moralnego użytego ewaluatywnie „Powinno się wykonać x” pociąga za sobą konieczność wykonania x. (Posiadanie moralnego motywu jest warunkiem wystarczającym działania z nim zgodnego).

Przekonanie o swojej akceptacji MI Hare ugruntował w czytelnikach *The Language of Morals* również za sprawą odpowiedzi, której udzielił na zarzut głoszący, że nie zawsze działamy zgodnie z akceptowaną przez nas powinnością. Podstawowym bowiem argumentem przeciwników internalizmu jest twierdzenie, że są sytuacje, w których akceptacja powinności nie motywuje ludzi do działania. Powyższa teza zmusza nas do przyjęcia poglądu, że nieprawdziwa jest albo mocna wersja internalizmu (MI), albo podane przez Ha-

²⁴⁶ Zob. tenże, *Internalism...*, dz. cyt., s. 97, 100.

²⁴⁷ Tenże, *The Language...*, dz. cyt., s. 20.

²⁴⁸ Zob. D. Brink, *Moral Realism and the Foundations of Ethics*, Cambridge University Press, New York 1989, s. 42; R. C. S. Walker, *Imperatywy kategoryczne*, w: T. Szubka (red.), *Metafizyka w filozofii analitycznej*, TN KUL, Lublin 1995, s. 522; W. Jacórzynski, *Hare Richard Mervyn. Moral thinking*, w: B. Skarga (red.), *Przewodnik po literaturze filozoficznej XX wieku*, tom 3, PWN, Warszawa 1995, s. 188; M. Rutkowski, *Obowiązek...*, dz. cyt., s. 265; A. MacIntyre, *Krótko...*, dz. cyt., s. 325.

re'a kryterium ewaluatywności sądów (KES). Hare z powyższego zarzutu wybrnął, ograniczając dziedzinę zastosowania KES. Jak stwierdził w *The Language of Morals*, nie każde użycie zwrotu powinnościowego musi pociągać za sobą imperatyw, gdyż sąd „Powinieniem zrobić X” może być traktowany jako złożenie trzech następujących sądów, z których tylko akceptacja trzeciego dostarcza motywu²⁴⁹:

(1) „Aby spełnić standardy, które zazwyczaj przyjmowane są przez ludzi, wymaga się wykonania X” (sąd odnoszący się do faktu socjologicznego).

(2) „Mam poczucie, że powinienem wykonać X” (sąd odnoszący się do faktu psychologicznego)

(3) „Powinieniem wykonać X” (sąd wartościujący).

Pomiędzy powyższymi trzema sensami zachodzi różnica znaczeniowa, ponieważ istnieją sytuacje, w których zastanawiamy się, czy nie wypowiedzieć, np. sądu wartościującego (3), który byłby niezgodny z (1) lub (2). Uświadamiają nam one, iż znaczenie „powinieniem wykonać x” jest złożone²⁵⁰. Choć często się zdarza, iż nasza wypowiedź moralna ma zarazem charakter (1), (2) i (3), to niekiedy owe sensory występować mogą oddzielnie. „Przypadki, w których sądy są postrzegane jako wartościujące, a nie pociągają za sobą imperatywów, zawsze po ich zbadaniu okazują się być przypadkami, w których sądy nie mają znaczenia (3), ale (1) lub (2) lub ich połączenie”²⁵¹. Hare wyraził to również inaczej: niemotywowanie do działania jest spowodowane użyciem „powinieniem zrobić x” w znaczeniu: (a) cudzysłowowym, (b) ironicznym lub (c) quasi-cudzysłowowym²⁵². W pierwszym przypadku nie dokonujemy autentycznego osądu moralnego, lecz cytujemy oceny innych, w drugim naszą intencją nie jest szczerze wypowiedzenie oceny, ale np. jej zanegowanie, w trzecim naszą intencją nie jest wyrażenie zalecenia, lecz stwierdzenie, iż dany czyn czy przedmiot posiada własności, które uznawane są przez społeczeństwo za rację danej oceny (użycie socjologiczne).

²⁴⁹ Por. R. M. Hare, *The Language...*, dz. cyt., s. 167.

²⁵⁰ Zob. tamże, s. 168.

²⁵¹ Zob. tamże.

²⁵² Zob. tamże, s. 124-126.

§ 4.13.1 Słabość woli

Ponieważ zawężenie dziedziny, w której obowiązuje kryterium ewaluatywności sądów (KES), potraktowano jako obronę internalizmu mocnego, wzmocniło to przekonanie, iż Hare akceptuje mocną wersję internalizmu (MI), co spotkało się z krytyką. Po pierwsze, podając w wątpliwość sposób ograniczenia dziedziny obowiązywania kryterium ewaluatywności sądów (KES), Hare'owi zarzucono arbitralność²⁵³. Można bowiem wątpić w wartość klasyfikacji różnych znaczeń zwrotów powinnościowych (psychologicznych, socjologicznych, ewaluatywnych itd.), jeśli nie istnieją ściśle kryteria takich podziałów²⁵⁴. Czy nie jest tak, że R. M. Hare uczynił to jedynie w celu obrony kryterium ewaluatywności sądów moralnych oraz mocnego internalizmu (MI)? Czy nie arbitralnie wprowadził on definicję sądu moralnego jako tego, który motywuje²⁵⁵, a następnie, spostrzegłszy jej nieadekwatność, arbitralnie zawęził dziedzinę jej zastosowania? Nie jest również prawdą, że jeśli wypowiedź „myślę, że powinienem zrobić x” nie powoduje działania, nie może mieć ona pierwotnego charakteru ewaluatywnego²⁵⁶. Istnieją bowiem sytuacje, w których szczerze i autentycznie akceptujemy obowiązek i czynimy wbrew temu, do czego on zobowiązuje. Jak pisze Donald Davidson:

Czy nie jest nigdy tak, że wydadę niezmałcony, stanowczy osąd, iż – wzięwszy wszystkie okoliczności pod uwagę – moje działanie nie jest najlepsze, a jednocześnie w postępku, którego się dopuszczam, nie ma śladu przymusu bądź

²⁵³ Zob. A. MacIntyre, *Krótko...*, dz. cyt., s. 327; M. Warnock, *Ethics...*, dz. cyt., s. 128-129; S. Lukas, *Moral Weakness*, w: G. Mortimore (ed.), *Weakness of Will*, Macmillan and St Martin's Press, London, Basingstoke 1971, s. 151.

²⁵⁴ Zob. M. Rutkowski, *Obowiązek...*, dz. cyt., s. 334-335.

²⁵⁵ Choć R. M. Hare często wypowiadał się tak, jakby prezentował definicje sprawodawcze, podany przez niego sposób definiowania ewaluatywności (KES) jest jedynie hipotezą. Jak pisze na temat KES w *The Language of Morals*: „Nie twierdzą w tym miejscu, iż dowodzę czegokolwiek materialnego na temat sposobu, w jaki używamy języka. Sugeruję jedynie terminologię, która jeśli zostanie zastosowana podczas badania języka moralnego, będzie, o czym jestem przekonany, udowadniała swoją wartość”; R. M. Hare, *The Language...*, dz. cyt., s. 168.

²⁵⁶ Zob. Rutkowski, *Obowiązek...*, dz. cyt., s. 335; H. Steward, *Akrasia*, w: E. Craig (ed.), *Routledge Encyclopedia of Philosophy*, Vol. 1, London and New York 1998, s. 139-142; A. Roskies, *Are Ethical Judgments Intrinsically Motivational? Lessons from „Acquired Sociopathy”, „Philosophical Psychology”* (16) 2003.

przymusowości? Nie ma dowodu na to, że takie sytuacje się zdarzają; lecz mnie wydaje się rzeczą absolutnie niewątpliwą, że się zdarzają²⁵⁷.

Są to przypadki tzw. słabości woli. O słabej woli (akrazji) mówimy wówczas, gdy ktoś autentycznie uważa, iż powinien coś zrobić – „całym sobą” podpisuje się pod autentycznym sądem moralnym, jednak zdarza się, że czyni rzecz przeciwną. Paradoksalność, dramatyczność i realność akrazji najlepiej oddaje znany fragment z Listu św. Pawła Apostoła do Rzymian:

Wiemy przecież, że Prawo jest duchowe. A ja jestem cielesny, zaprzędany w niewolę grzechu. Nie rozumiem bowiem tego, co czynię, bo nie czynię tego, co chcę, ale to, czego nienawidzę – to właśnie czynię. Jeśli zaś czynię to, czego nie chcę, to tym samym przyznaję Prawu, że jest dobre²⁵⁸.

Podane powyżej ograniczenia obowiązywania kryterium ewaluatywności sądów oraz internalizmu mocnego zaprzeczają jednak istnieniu takich sytuacji. Gdyby te zasady były prawdziwe, należałoby powiedzieć, że autentyczna akrazja jest jedynie epifenomenem – akceptacja autentycznej powinności zawsze pociąga za sobą jej wykonanie, a jeśli tak się nie dzieje, nie była to autentyczna akceptacja bądź powinność. Nasza intuicja lingwistyczna wskazuje jednak, że nie jest to prawdą – błędna jest raczej teoria Hare'a. Kolejną konsekwencją jednoczesnego obowiązywania mocnej wersji internalizmu oraz kryterium ewaluatywności sądów jest niemożność niesprecznego wypowiedzenia autentycznego sądu moralnego, który ma np. postać: „Mogę zrobić to, o czym myślę, że nie powinienem”. Jeśli autentyczna słabość woli nie mogłaby istnieć na mocy praw pojęciowych, nie można byłoby sensownie wypowiedzieć powyższego zdania²⁵⁹.

Jednak, jak wskazaliśmy powyżej, Hare nie był zwolennikiem mocnego internalizmu motywacyjnego, choć prawdą jest, że zakładał kryterium ewaluatywności sądów moralnych (KES). Już we *Freedom and Reason* odpowiedział na część powyższych zarzutów oraz przestał formułować tezy wskazujące na akceptację MI. Nigdy jed-

²⁵⁷ D. Davidson, *Jak jest możliwa słabość woli*, w: J. Hołówka (red.), *Filozofia moralności. Postanowienie i odpowiedzialność*, Spacja – Fundacja Aletheia, Warszawa 1997, s. 90.

²⁵⁸ List do Rzymian, 7, 14-17.

²⁵⁹ Zob. M. C. McGuire, *Can I Do What I Ought Not? Where Has Hare Gone Wrong?*, „Mind” (279) 1961, s. 400; D. Brink, *Moral Realism...*, dz. cyt., s. 48; W. Jacyński, *Hare...*, dz. cyt., s. 188; M. Rutkowski, *Obowiązek...*, dz. cyt., s. 335.

nak nie przyznał, że jego poglądy w kwestii motywacji uległy zmianie²⁶⁰. Stwierdził jedynie, że we wczesnych swych pracach nie sformułował ich wystarczająco jasno. Uważał, iż we *Freedom and Reason* nie zmienił swoich poglądów, tylko wyjaśnił te elementy *The Language of Morals*, które nie były klarowne i powodowały błędną interpretację jego myśli²⁶¹. Po pierwsze, odciął się od poglądu, iż istnieją różne znaczenia zwrotu „powiniennem zrobić x”²⁶². Znaczy on w każdym kontekście to samo, choć w różnych użyciach wywołuje inne postawy i stany umysłu. Po drugie, nieznacznie zmienił podejście również do problemu akrazji. Nie starał się wyjaśnić jej poprzez wykluczanie takich sensów sądów moralnych, które nie powodują działania – jak uczynił to w *The Language of Morals* – lecz wskazywał na konkretne stany fizyczne lub umysłowe, które powodują brak działania²⁶³. Problem akrazji przesunął z płaszczyzny semantycznej na psychologiczną – wyjaśnienia semantyczne zastąpił psychologicznymi. Według niego, nasze stany umysłu, korespondujące z danym sądem powinnościowym, mogą ulec zmianie, choć jego znaczenie pozostanie niezmiennie. We *Freedom and Reason* stwierdził, że zwrot powinnościowy może utracić swą motywującą siłę bez zmiany swego znaczenia na wiele sposobów²⁶⁴. Po trzecie, w pracy tej dodał, że w *The Lanugage of Morals* nie było jego celem zaprzeczenie istnienia autentycznej akrazji, a jedynie podanie takich przypadków, w których przyczyną braku działania nie jest autentyczna słabość woli. Często bowiem nieprzestrzeganie powinności wyjaśniamy rzekomą słabością woli, choć zostały niedopełnione jedynie pewne potrzebne do działania warunki. W związku z tym w kolejnych tekstach prezentował coraz bogatszy wachlarz postaw, mylnie postrzeganych jako akrazję. We *Freedom and Reason* pisał o hipokryzji moralnej, schodzeniu

²⁶⁰ „W *The Language of Morals*, w bardzo wstępnym ujęciu problemu słabości woli, stwierdziłem w rzeczy samej, że w niektórych jej wypadkach chodziło o sądy, którym brakowało cech nakazu. I takie wypadki, rzecz jasna, się zdarzają. [...] Jednak przypuszczenie, iż wszystkie wypadki słabości woli są takiego właśnie rodzaju, świadczyłoby o braku przenikliwości, toteż nawet wówczas tak nie myślałem”; tenże, *Myślenie...*, dz. cyt., s. 77.

²⁶¹ Zob. R. M. Hare, *Freedom...*, dz. cyt., s. 79, 84.

²⁶² Zob. tamże, s. 75-76.

²⁶³ Zob. G. Mortimore, *Editor's Introduction*, w: G. Mortimore (ed.), *Weakness of Will*, Macmillan, St Martin's Press, London and Basingstoke 1971, s. 14.

²⁶⁴ Zob. R. M. Hare, *Freedom...*, dz. cyt., s. 75-76.

na złą drogę, nieszczerości, niemocy fizycznej lub psychicznej²⁶⁵. Jednak najbardziej wyczerpujące rozróżnienie takich przypadków znaleźć można w artykule *Weakness of Will*²⁶⁶. Hare podał tam następujące wyjaśnienie ludzkich działań uważanych za sprzeczne z internalizmem²⁶⁷:

Wyjaśnienie działań akratycznych

1. Ktoś może akceptować zasadę moralną, lecz błędnie uważać, że nie stosuje się ona do danej sytuacji (niedoskonała uniwersalizacja).
2. Ktoś może samooszukiwać się lub nie być szczerzy: mówi, że powinno się coś zrobić, lecz wcale tak nie myśli (nieautentyczna akceptacja zasady).
3. Ktoś może wypowiadać jakąś zasadę, ale jej nie rozumieć (np. człowiek pijany).
4. Ktoś może uważać że powinien coś zrobić, lecz nie może tego uczynić. W grę może wchodzić:
 - a. niemożność fizyczna;
 - b. niemożność psychiczna (niemożność oparcia się przeciwstawnym pokusom).
5. Ktoś może uważać, że dany czyn jest wymagany przez społeczność, lecz nie akceptować tego przekonania i nie traktować go jako nakazu. Taka osoba używa sądu moralnego nie w sensie preskryptywnym, lecz „cudzysłowowym” (socjologicznym).
6. Ktoś może myśleć, że powinien coś zrobić, myśli jednak w słabym sensie słowa „myśleć”:
 - a. posiada uczucie zobligowania do zrobienia tego, a nie mocną akceptację danego nakazu (używa sądu moralnego w sensie psychologicznym);
 - b. ktoś myśli, że powinien, ale jego „myślenie” nie jest wystarczająco pewne, aby działać zgodnie z tym sądem (akceptacja nie ma wystarczającej mocy).
7. Pewna „część kogoś” uważa, że powinien coś zrobić z całą preskryptywną mocą, lecz „druga jego część” opiera się i uniemożliwia mu zrobienie tego. Owa „druga część” to ogarniająca go w dosłownym sensie słabość woli.

²⁶⁵ Zob. tamże, s. 67-86. Por. S. Lukas, *Moral...*, dz. cyt., s. 149-150.

²⁶⁶ R. M. Hare, *Weakness of the Will*, w: tenże, *Objective Prescriptions...*, dz. cyt.

²⁶⁷ Zob. tamże, s. 112-113.

W powyższym *Wyjaśnieniu działań akratycznych* zauważyć można trzy rodzaje podejścia do problemu akrazji. Po pierwsze, w *Weakness of Will* Hare potwierdził możliwość wyjaśnienia niektórych sytuacji braku działania w taki sposób, iż stają się one prawdami analitycznymi. Wzbogacił również sposób takiego wyjaśnienia. Powołując się na wprowadzony w *Myśleniu moralnym* podział ocen na intuicyjne i krytyczne²⁶⁸ (płaszczyznę myślenia intuicyjnego i krytycznego), stwierdził, że częstym powodem występowania braku związku motywacyjnego jest fakt, że akceptowane przez nas zasady moralne mają źródło w myśleniu intuicyjnym. Ono natomiast jest jedynie quasi-moralne; intuicje moralne są psychologicznymi lub socjologicznymi stwierdzeniami. Zgodził się więc z poglądem wyrażonym w *The Language of Morals*, że nie „wszystkie sądy powinnościowe pociągają za sobą imperatywy – czynią tak tylko te z nich, które są używane oceniająco”²⁶⁹. Choć moc illokucyjna intuicyjnych zasad moralnych także jest preskryptywna, ich nakazowość, jak pisał Hare:

jest tak mocno zdominowana przez przekonanie o ich prawdziwości, że czasami osoba uznająca ten sąd potrafi przypisać mu prawdziwość pomimo tego, iż nie jest nastawiona na jego przestrzeganie i tego nie robi²⁷⁰.

Przykłady zaniku normatywności, charakterystyczne dla zasad moralnych używanych na płaszczyźnie intuicyjnej, opisują punkty 2, 5 i 6a przedstawionego powyżej *Wyjaśnienia działań akratycznych*.

Po drugie, Hare stwierdził, tak jak we *Freedom and Reason*, że brak działania może wynikać z ułomności człowieka – błędów poznawczych, niedoskonałej akceptacji, niewystarczającej pewności co do przekonań moralnych czy powierzchowności uniwersalizacji. Ludzie często myślą się co do faktów, ulegają słabości intelektualnej, brak im wyobraźni, wrażliwości czy konsekwencji w myśleniu. Często uważają, że wydali sąd moralny, ale niedoskonałość ich aparatu poznawczego doprowadza ich jedynie do sądu quasi-moralnego. Takie przypadki zostały opisane w punktach 1, 3 oraz 4a *Wyjaśnienia działań akratycznych*.

Po trzecie, potwierdził, iż istnieją autentyczne sytuacje słabości woli, które nie mogą być wyjaśnione w żaden z powyższych sposo-

²⁶⁸ Zob. § 6.1.

²⁶⁹ Tenże, *The Language...*, dz. cyt., s. 164.

²⁷⁰ Tenże, *Weakness...*, dz. cyt., s. 113.

bów. Fenomeny 4b, 6b, 7 *Wyjaśnienia działań akratycznych* wskazują, iż problem akrazji jest nierozwiązywalny za pomocą analizy pojęciowej. Jest to problem psychologiczny – autonomiczny względem semantyki czy logiki moralności²⁷¹. Akrata w tej interpretacji to człowiek, który akceptując zasady moralności, nie ma siły czy psychicznej możliwości, aby ich przestrzegać. Ludzie mają bowiem słabą wolę – kierują się często arbitralnymi pobudkami, które są sprzeczne z logiką. Z tego powodu Hare w punktach 4b, 6b, 7 *Wyjaśnienia działań akratycznych* odwołuje się do wyjaśnień psychologicznych. Sytuacja słabości woli, opisana przez św. Pawła, jest psychologiczną niemocą działania zgodnego z tym, co słuszne i dobre. Nawet jeśli akrata szczerze uważa, że ma obowiązek oraz to przekonanie spełnia wszystkie formalne warunki, tj. jest ono mocną, szczerą akceptacją uniwersalnego zalecenia o charakterze nadrzędności, to i tak mogą zdarzyć się sytuacje, w których nie podejmie on działania. Takie przypadki najtrafniej opisują dwie metafory: krnąbrności naszej natury („ciała”) lub wewnętrznego rozdarcia (jedna „część nas” nakazuje wykonać prawo, lecz druga się temu sprzeciwia). Te dwie metafory są tak naturalne dla naszego języka potocznego, że filozofowie, którzy próbują uściślić ich znaczenie, muszą wymyślić swój własny język²⁷². Intencją Hare’a nie było więc zredukowanie problemu akrazji przez stwierdzenie, że jeśli używamy języka oceniająco, akrazja nie jest możliwa, a jeśli nie używamy go oceniająco, nie może być mowy o autentycznej słabości woli. Już we *Freedom and Reason* uważał on bowiem, iż problem słabości woli jest realny i potrafimy go wyjaśnić jedynie za pomocą metafor odwołujących się do niemocy psychicznej. Co więcej, paradoks słabej woli świadczy na rzecz przedstawionych praw internalizmu moralnego (PIM1–PIM4). Jeśli bowiem akrazja jawi nam się jako praktyczna anomalia, a w konsekwencji problem teoretyczny, świadczy to o tym, iż nasze intuicje dotyczące języka moralnego zakładają prawdziwość internalizmu motywacyjnego. Gdyby rację mieli eksternaliści motywacyjni, nie czuliśmyby paradoksalności przypadku opisanego przez Pawła z Tarsu. To właśnie prawdziwość internalizmu powoduje zagadkowość słabo-

²⁷¹ Chociaż Hare nie przedstawił takiej tezy *explicite*, wydaje się, że jest ona główną myślą eseju *Weakness of the Will*. Zob. również: tenże, *The Language...*, dz. cyt., s. 169.

²⁷² Zob. tenże, *Freedom...*, dz. cyt., s. 81.

ści woli²⁷³. Argument ten jest jednak słaby. Twierdzenie, że „akrazja potwierdza autentyczność internalizmu”, może wydać się tak przewrotne jak głoszenie, że wyjątek potwierdza regułę. Wyjątki obalają reguły, a na pewno uderzają w konieczny ich charakter.

Analiza pojęciowa, logika moralności doprowadziła Hare’a od formalnego prawa internalizmu semantycznego (PIM3) do myślenia krytycznego – metody, którą czasami musimy stosować, jeśli chcemy dokonywać poprawnych ocen²⁷⁴. Internalizm traktował jako zbiór praw czysto pojęciowych. Jego zdaniem to stanowisko wypływa z analizy językowej, ze sposobu użycia słów, a nie z założeń psychologicznych na temat ludzi. Twierdził:

Bronię jedynie pojęciowego połączenia pomiędzy *myśleniem*, że powinno się coś zrobić, a byciem motywowanym do zrobienia tego, oraz pomiędzy *myśleniem*, że ktoś inny powinien coś zrobić, a chceniem, aby to zrobić²⁷⁵.

Zawarty tu pogląd będziemy nazywać **internalizmem pojęciowym**. Tego typu teza nie pociąga za sobą internalizmu mocnego – nie mówi niczego na temat działania, a jedynie na temat wynikania imperatywów (ewentualnie dyspozycyjnej motywacji lub konieczności chcenia). Fakt obowiązywania pewnej logiki moralnej gry językowej – logiki moralności – nie determinuje tego, że zawsze podporządkujemy jej nasze działanie, choć stanowi ona dla nas wzór, do którego dążymy. Podobna zależność zachodzi w dziedzinie sądów opisowych: to, że obowiązują konieczne reguły wynikania opisane przez logikę klasyczną, nie znaczy, że z konieczności wszyscy faktycznie się do nich stosują. To, że oszukujemy, grając w brydża, nie znaczy, że reguły gry w brydża nie istnieją lub jest w nich zapisana możliwość oszukiwania. Analogiczna zasada dotyczy również moralności. Prawa internalizmu moralnego (PIM1–PIM4), a nade wszystko mocna wersja internalizmu (MI), które Hare próbował wyprowadzić z logiki moralności, są jedynie wzorem etycznej racjonalności. Człowiek może jedynie próbować działać racjonalnie. „Grając w moralność”, chcielibyśmy grać zgodnie z jej regułami, często jednak łamiemy jej zasady. Związek pomiędzy szczerym powiedzeniem „zrobię x” a działaniem nie jest więc konieczny – zasada internalizmu mocnego (MI) jest nieprawdziwa. Z tego powodu zdanie: „Mogę zrobić to, o czym

²⁷³ Zob. tamże, s. 69.

²⁷⁴ Na temat metody myślenia krytycznego zobacz § 6.

²⁷⁵ Tenże, *Internalism...*, dz. cyt., s. 98.

myślę, że nie powinienem” w świetle teorii Hare’a, jak sugerował M. C. McGuire²⁷⁶, nie jest wewnętrznie sprzeczne. Zdania takiego zgodnie z prawdą nie mógłby wypowiedzieć jedynie Hare’owski idealny nakazodawca – archanioł, który został opisany w kolejnej części rozprawy. Gdy wygłasza je człowiek, znaczy ono jedynie:

Mam słabą wolę, czasem nie działam względem racji, ale z nieracjonalnych pobudek, nie zawsze również przestrzegam logiki, zatem choć uważam, że czyn x jest zły, może się stać tak, iż go wykonam.

Można pokusić się o tezę, że motyw powoduje działanie dla archanioła, dla ludzi natomiast je tylko „konstruuje”. M. Rutkowski motywy, które jedynie „konstruuja” działanie, nazywa motywami *dyspozycyjnymi*, przeciwstawiając je motywom *faktycznym*²⁷⁷. Przychylając się do tego nazewnictwa, w konsekwencji można byłoby powiedzieć, iż internalizm Hare’a jest internalizmem słabym – tylko archanioł bowiem mógłby być „bohaterem” internalizmu mocnego. Musielibyśmy tak stwierdzić, gdybyśmy internalizm mocny nazwali teorią, w ramach której obowiązek nierozzerwalnie łączy się z działaniem, internalizm słaby zaś koncepcją, zawierającą *dyspozycyjne* znaczenie motywu²⁷⁸.

Biorąc pod uwagę to, co przedstawiono powyżej, internalizm R. M. Hare’a można zobrazować w następujący sposób:



Rysunek 6. Internalizm R. M. Hare’a

²⁷⁶ Zob. M. C. McGuire, *Can...*, dz. cyt., s. 400.

²⁷⁷ Zob. M. Rutkowski, *Obowiązek...*, dz. cyt., s. 37, 148-149.

²⁷⁸ David Brink internalizm mocny i słaby definiuje w następujący sposób: „Słaby internalizm motywacyjny głosi, że jest prawdą pojęciową, iż rozważania moralne dostarczają *pewnej* motywacji, podczas gdy mocny internalizm motywacyjny twierdzi, że jest prawdą pojęciową, iż rozważania moralne dostarczają *wystarczającą* motywację do działania”; D. Brink, *Moral Realism...*, dz. cyt., s. 41.

Powyższą interpretację uzasadnia również przekonanie Hare'a na temat charakteru samej motywacji. Biorąc pod uwagę płaszczyzną psychologiczną, należy stwierdzić, że nie zgadzał się on z emotywistami, którzy uważają, iż funkcja motywacyjna może być wyjaśniona poprzez kauzalne oddziaływanie wypowiedzi na nasze uczucia – że zalecenia motywują dlatego, iż wzbudzają emocje. Preskrypcje nie są ekspresją naszych życzeń czy uczuć, ale racjonalnej woli²⁷⁹. Według niego, emotywiści popełnili błąd, pomieszczyli bowiem płaszczyznę semantyczną z psychologiczną i w konsekwencji ich teoria jest psychologiczna, nie zaś semantyczno-logiczna²⁸⁰. Sądy moralne nie zmieniają naszych postaw przez perswazję, lecz motywują nas dzięki racjom, bowiem, tak jak wszystkie zalecenia, podlegają prawom logiki²⁸¹. Nie są one również z konieczności wyrazem naszych uczuć czy życzeń²⁸². Istotne logiczne własności sądów moralnych – ich preskryptywność, uniwersalność i nadrzędność – są w nich obecne nawet wówczas, gdy ich użycie nie jest związane z żadnymi emocjami²⁸³.

Ponieważ prawa logiki języka moralnego – które dla Hare'a są prawami pojęciowymi²⁸⁴ – nie mają charakteru psychologicznego czy przyczynowo-skutkowego, również internalizm, który jest ich pochodną, nie może mieć takiej natury. Hare nie popierał żadnej z wersji **Hume'owskiego modelu motywacyjnego**, którego główną przesłanką jest „reguła instrumentalnych racji praktycznych: mając rację do działania, ma się też pragnienia”²⁸⁵. Nie zgodziłby się on z tezą, iż rozum nie może przeciwstawić się pragnieniom w kierowaniu zachowaniem, że jest „niewolnikiem uczuć”. Podzielał raczej pogląd Arystotelesa, który, jak pisze Hare,

w słynnej metaforze [...] powiada, że umysł do charakteru [...] musi mieć stosunek rodzicielski: dlatego że motywy i dyspozycje człowieka są racjonalne, i z tego właśnie powodu „słuchają rozumu jak ojca”²⁸⁶.

²⁷⁹ R. M. Hare, *Prescriptivism*, w: E. Craig (ed.), *Routledge Encyclopedia of Philosophy*, Vol. 7, Routledge, London and New York 1998, s. 669.

²⁸⁰ Zob. tenże, *Sorting...*, dz. cyt., s. 15-18.

²⁸¹ Zob. tenże, *The Language...*, dz. cyt., s. 13.

²⁸² Zob. tamże, s. 12.

²⁸³ Tamże, s. 144.

²⁸⁴ Zob. tenże, *Ontology...*, dz. cyt., s. 97.

²⁸⁵ M. Rutkowski, *Obowiązek...*, dz. cyt., s. 71.

²⁸⁶ R. M. Hare, *Myslenie...*, dz. cyt., s. 63.

Powolywał się on również na Kanta: swój internalizm prezentował jako tezę głoszącą, że jeśli ktoś szczerze zaakceptuje daną zasadę moralną, musi chcieć („w szerokim sensie tego słowa, które zawiera kantowską *wolę* oraz trzy rodzaje arystotelesowskiego *orexis*”²⁸⁷) działać zgodnie z tą zasadą. W tym świetle przypisywanie Hare'owi akceptacji mocnej wersji internalizmu (MI) jako zasady rządzącej psychiką człowieka wydaje się interpretacją nietrafną. Zakładanie, że rozum zawsze doprowadza do działania, że z konieczności jest niezłomnym kierownikiem naszych działań, byłoby bowiem ewidentnie niezgodne z codziennym doświadczeniem wszystkich bodaj ludzi oraz tradycją filozoficzną, która od czasów Davida Hume'a boryka się z przeciwnym problemem: jak uzasadnić choćby minimalną motywacyjną rolę rozumu.

Z tego również powodu zaliczenie internalizmu Hare'a do tzw. internalizmu postawy jest posunięciem nietrafnym²⁸⁸. **Internalizm postawy** to stanowisko, które głosi, że osoba aprobująca obowiązek moralny ma motyw do działania, ponieważ wykazuje w stosunku do niego postawę aprobaty (*pro-attitude*). Taki pogląd akceptowali niewątpliwie emotywiści, jednak nie R. M. Hare. Owszem, zgodziłby się on, że akceptacja zalecenia powodowana jest niekiedy poczuciem aprobaty w stosunku do danego działania; że zalecenia, które wygłaszamy, często motywują nas dlatego, że są wydawane pod wpływem uczuć. Możemy jednak szczerze zaakceptować i wykonać rozkaz, nie darząc działania, które mamy przedsięwziąć, żadnym pozytywnym uczuciem – co więcej, żywiąc w stosunku do niego odrazę lub trwogę. Możemy również zaniechać działania, co do którego mamy uczucie

²⁸⁷ Tenże, *Internalism...*, dz. cyt., s. 96-97; tenże, *Prescriptivism*, w: tenże, *Objective Prescriptions...*, dz. cyt., s. 21; tenże, *Prescriptivism*, w: red. E. Craig, *Routledge Encyclopedia of Philosophy*, Vol. 7, Routledge, London and New York 1998, s. 669.

²⁸⁸ Zob. M. Rutkowski, *Obowiązek...*, dz. cyt., s. 270-271. Oczywiście to, czy stanowisko Hare'a należy nazwać internalizmem postawy, zależy od tego, co rozumiemy przez pojęcie „postawy”. Jeśli postawa aprobaty (*pro-attitude*) ma charakter uczuciowy, nie można nazywać Hare'a zwolennikiem internalizmu postawy (M. Rutkowski, krytykując internalizm postawy, mówi o „uczuciu aprobaty”). Z drugiej strony, jeśli przyjmiemy definicję postawy jako zjawiska przynależnego do sfery emocjonalno-wolitywnej, w których dochodzi do głosu przychylny lub nieprzychylny stosunek do rzeczywistości, a wśród takich wymieni się m.in. zamierzenia, aspiracje, potrzeby, preferencje, pragnienia itp. (por. M. Fritzhand, *Główne zagadnienia i kierunki metaetyki*, Książka i Wiedza, Warszawa 1970, s. 306), to teorię Hare'a można nazwać internalizmem postawy.

aprobaty. Rozdźwięk między uczuciem, pragnieniem a obowiązkiem jest jeszcze silniejszy, ponieważ sądy moralne są nie tylko zaleceniami, lecz zaleceniami zuniwersalizowanymi. To m.in. ta cecha sprawia, że można mieć obowiązek, nie będąc do niego przychylnie nastawionym²⁸⁹. Z drugiej jednak strony, choć Hare nie był internalistą postawy, niewątpliwie akceptował on **internalizm epistemiczny** (nie może nie mieć motywów do działania osoba w jakiś sposób aprobująca obowiązek moralny²⁹⁰). Prawa internalizmu moralnego PIM1–PIM4 opierają się bowiem na pojęciu akceptacji (*assent to*) czy podpisania się pod danym zdaniem (*subscribe to*), zawierają więc element epistemiczny. W świetle etyki Hare'a racja do działania jest szczerą akceptacją sądu moralnego, utożsamiał on bowiem rację praktyczną z „podpisaniem się” pod zasadą „wynikającą” z pewnych racji logicznych²⁹¹ – z akceptacją danej maksymy, z jej zaakceptowaniem jako racji do działania²⁹².

§ 4.13.2 Sceptyczny amoralista

Istnieje jeszcze jeden argument skierowany przeciwko internalizmowi Hare'a. Głosi on, iż w ramach przyjętych przez niego założeń nie można wyjaśnić zachowania sceptycznego amoralisty. Wskazywanie na trudności w uzgodnieniu internalizmu z istnieniem amoralistów stało się dyskutowanym argumentem za sprawą Davida Brinka²⁹³. Według niego, **sceptyczny amoralista** to człowiek, który rozpoznając pewne moralne zasady, nie jest przez nie motywowany. Postępuje on wbrew uznanemu przez siebie zbiorowi moralnych reguł rządzących zachowaniem. Nie jest to jednak osoba, która, jak często w przypadku słabości woli, działa „w afekcie” – ma „rozdartą duszę”, gdyż ulega „pokusie ciała”, „jedna jej część chce działać zgodnie z obowiązkiem,

²⁸⁹ Zob. § 5.2.

²⁹⁰ Por. M. Rutkowski, *Obowiązek...*, dz. cyt., s. 270-271.

²⁹¹ Racja logiczna stanowi uzasadnienie treści danego sądu moralnego, przy czym wyznaczona jest ona często przez standardy akceptacji danej zasady moralnej. Zob. R. M. Hare, *Internalism...*, dz. cyt., s. 102-103. Rozróżnienia na racje praktyczne i logiczne używamy za M. Rutkowskim. Racje logiczne nazywa on także „racjami co do przekonań”. Termin „racje” praktyczne będziemy traktowali jako synonim „racji do działania”. Zob. M. Rutkowski, *Obowiązek...*, dz. cyt., s. 54.

²⁹² Zob. R. M. Hare, *Universalisability*, w: tenże, *Essays on the Moral...*, dz. cyt., s. 15.

²⁹³ Zob. D. Brink, *Externalist Moral Realism*, „Southern Journal of Philosophy”, Supplement (24) 1989; zob. tenże, *Moral Realism...*, dz. cyt., s. 46-49.

lecz druga się temu sprzeciwia”. Zdaniem Brinka, możemy wyobrazić sobie kogoś, kto konsekwentnie, z własnej woli oraz świadomie uchyla się od czynienia tego, co uważa za wymóg moralności, nie używając wcale zwrotów moralnych w mowie zależnej, w znaczeniu cudzysłowowym, socjologicznym itp²⁹⁴. Amoralista często będzie więc osobą niemoralną – świadomie wybierającą moralne zło jako zło. Zdaniem Brinka, bez trudu możemy pomyśleć istnienie takiej osoby, co powinno stanowić problem, jeśli internalizm byłby prawdziwy.

Hare nie zgodził się jednak na możliwość istnienia powyżej opisanego amoralisty, według niego bowiem nikt racjonalny nie może świadomie i szczerze pragnąć tego, co uważa za zło. Osoba, która używa sądów moralnych w znaczeniu ewaluatywnym, tzn. zna formalne reguły nimi rządzące oraz szczerze je akceptuje, nie może nie posiadać choćby motywu dyspozycyjnego, chyba że zachodzi któryś z przypadków przedstawionych w *Wyjaśnieniu działań akratycznych*. Nie znaczy to jednak, że Hare negował możliwość istnienia sceptycznych amoralistów – nie uznawał jedynie takiego sposobu ich opisu, jaki zaprezentował D. Brink. Amoralizm Hare zinterpretował jako postawę kogoś, kto

powstrzymuje [wyróżnienie K. S.] się od wypowiedzania sądów moralnych [...] albo wypowiada jedynie takie, które są moralnie obojętne, i postępuje tak albo zawsze [amoralista konsekwentny – K. S.], albo w odniesieniu do konkretnych przypadków [amoralista wybiórczy – K. S.]²⁹⁵.

Jest to osoba, która uważa, iż pewne lub wszystkie działania są moralnie neutralne²⁹⁶. Twierdzi ona, że pewne działania są moralnie neutralne, jeśli np. powiedzenie, że powinny albo nie powinny zostać podjęte, jest sprzeczne z jej interesem lub innymi akceptowanymi przez nią praktykami oceniania²⁹⁷. Amoralista jest więc człowiekiem, który nie tyle jest niemoralny, co stawia siebie poza moralnością, nie uznając kwalifikacji moralnych. Według Hare'a, taka postawa, jeśli jest konsekwentna, nie musi być wewnętrznie sprzeczna. Co więcej, amoralista nie musi porzucać języka moralności²⁹⁸. Może

²⁹⁴ Zob. tenże, *Externalist...*, dz. cyt., s. 30; tenże, *Moral Realism...*, dz. cyt., s. 46.

²⁹⁵ R. M. Hare, *Myslenie...*, dz. cyt., s. 228.

²⁹⁶ Zob. tamże, s. 231.

²⁹⁷ Zob. tamże, s. 230.

²⁹⁸ Zob. tamże, s. 228. Rozwiązania, którego użył Hare w *The Language of Morals* (rozdzielenia na znaczenie ewaluatywne i cudzysłowowe), aby odnieść się do problemu akrazji, Michael Smith broni w zastosowaniu do problemu amo-

on doskonale znać taki język i wszystkie jego prawa oraz potrafić się nim posługiwać,

ale słów tworzących dyskurs moralny [...] używać w zdaniach rozpoczynających się w taki oto sposób: „Nie chodzi o to, że” [«...nie powinienem, ale o to, że chcę» – K. S.]. Przypomina kogoś, kto wie, jak używać słowa „czarownica”, ale w istnienie czarownic nie wierzy; wie, jakie mogą być obowiązki, ale [...] nie dopuszcza istnienia jakichkolwiek²⁹⁹.

Jedynie amoralista wybiórczy, który raz używa języka moralnego, a innym razem nie przestrzega jego formalnych własności, jest wewnętrznie sprzeczny³⁰⁰. Biorąc bowiem pod uwagę to, że sądy moralne z konieczności muszą być uniwersalne, nie mógłby on spójnie i szczerze raz powoływać się na powinności i obowiązki, a przy innej okazji siebie z nich wykluczać. Nie można zuniwersalizować maksymy: „Przestrzegaj reguł moralnych i wszystkich praw moralnej gry językowej, jeśli jest to w twoim interesie lub gdy masz na to ochotę” – nikt bowiem nie mógłby chcieć³⁰¹, aby obowiązywało takie powszechne prawo. Można jednak konsekwentnie działać tak, jakby wszystkie czyny były moralnie neutralne³⁰².

W świetle etyki Hare’a nie musimy zakładać, że amoralista konsekwentny jest osobą nieracjonalną, gdyż taka postawa nie musi być wewnętrznie sprzeczna³⁰³. Zdaniem Hare’a, osoby posługujące się sądami moralnymi muszą to czynić zgodnie ze wszystkimi prawami pojęciowymi języka moralnego pod zarzutem sprzeczności logicznej lub błędnego uchwytowania znaczeń, jednak jeśli ktoś konsekwentnie nie chce używać języka moralnego, może tego nie czynić,

ralisty. Według niego sceptyczny amoralista to człowiek, który potrafiąc używać terminów i sądów moralnych, nie chwyta w pełni ich znaczenia. Aby uzasadnić, iż nie jest to rozwiązanie *ad hoc*, odwołał się on do analogii: tak jak uchwycenie znaczenia pojęcia „czerwony” wymaga posiadania pewnego stanu psychicznego – doświadczenia czerwoności, tak samo uchwycenie znaczenia pojęć moralnych wymaga posiadania innego stanu psychicznego – tj. motywacji do działania. Brink niesłusznie więc zakłada, że amoralisci trafnie uchwytują znaczenia sądów moralnych, nawet jeśli nie posiadają motywacji. Według Smitha teza internalizmu jest mocniej uzasadniona niż przekonanie, że amoralisci spełniają powyższy warunek. Zob. M. Smith, *The Moral...*, dz. cyt., s. 66-71.

²⁹⁹ R. M. Hare, *Myślenie...*, dz. cyt., s. 229.

³⁰⁰ Zob. tamże, s. 230-232.

³⁰¹ W świetle filozofii Hare’a należy mówić o logicznej niemożności chcenia oraz niemożności praktycznej (nie byłoby to w naszym interesie). Zob. § 6.2.

³⁰² Zob. tamże, s. 232.

³⁰³ Zob. tamże, s. 238; A. MacIntyre, *Krótko...*, dz. cyt., s. 19-20.

nie obawiając się sprzeczności. Z drugiej jednak strony, choć powyższy sposób wyjaśnienia postawy amoralisty neutralizuje zarzut D. Brinka, znacznie osłabia on pozycję uniwersalnego preskrytywizmu.

Zdaniem A. MacIntyre’a³⁰⁴ to, iż „Hare dopuszcza możliwość bycia amoralistą bez popadania w irracjonalizm, [...] implikuje [...] że u podłoża wszelkich sądów moralnych znajduje się raczej nieracjonalny, przedracjonalny wybór moralności aniżeli stan amoralizmu. A zatem to, co zrazu zdawało się stanowiskiem będącym bardziej adekwatnym przeformułowaniem Kantowskiego utożsamienia wymogów moralności z kategorycznymi wymogami rozumu praktycznego, okazuje się – wbrew intencjom i twierdzeniom Hare’a – bliższe w swych niezamierzonych implikacjach Kierkegaardowskiemu *albo-albo*”³⁰⁵. Powyższa konkluzja wydaje się prawdziwa. Również Hare traktował ją jako słabość swojej teorii³⁰⁶. Według niego,

logika nie może prowadzić nas przez cały czas od wiary w fakty pozamoralne do dyspozycji do działania. Możemy więc wybrać albo rozziw między przekonaniem dotyczącym faktów a sądami moralnymi³⁰⁷, tak jak to robią preskryptywiści, albo, co wolą deskryptywiści, rozziw między sądami moralnymi a dyspozycjami do działania³⁰⁸, utrzymując ścisły związek między przekonaniem dotyczącym faktów a sądami moralnymi. [...] Ponieważ nie jesteśmy deskrytywistami, nie możemy utrzymywać, że istnieje ścisły związek między przekonaniem dotyczącym faktów a sądami moralnymi; to zaś prowadzi, jak widzieliśmy, do tego, że konsekwentnemu amoralistcie pozostawia się drogę ucieczki: może uznać wszystkie nasze fakty, a mimo to nie formułować żadnych sądów moralnych³⁰⁹.

Hare próbował jednak, wbrew ocenie MacIntyre’a, uzasadnić pogląd, że warto być moralnym. Nie przedstawił on jednak racji kategorycznych, z którymi każdy rozumny podmiot musi się zgodzić – dowodu wyprowadzonego z analizy znaczeń, logiki czy czystego rozumu. Odwoływał się jedynie do empirycznych przekonań na temat psychiki ludzi, wskazywał na związane z amoralizmem niebezpieczeństwa, starając się wykazać, iż rozważnie (zgodnie z własnymi

³⁰⁴ Podobny zarzut znaleźć można w: T. Nagel, *Foundations...*, dz. cyt., s. 12. Omawia go również Peter Singer w: tenże, *Reasoning towards Utilitarianism*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 147-159.

³⁰⁵ A. MacIntyre, *Krótko...*, dz. cyt., s. 19-20.

³⁰⁶ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 232.

³⁰⁷ Zob. § 3.16.

³⁰⁸ Zob. § 4.3.

³⁰⁹ Tamże, s. 237.

przekonaniami oraz we własnym, egoistycznym interesie³¹⁰) jest być osobą moralną. Nie jest jednak tak, iż wszystkie zasady wywieść można z egoistycznej rozważki³¹¹, choć z utylitarystycznego punktu widzenia – dla społeczeństwa – rozważne jest przestrzeganie zasad moralnych³¹². Argumenty Hare’a przedstawione w jedenastym rozdziale *Myślenia moralnego* nie dowodzą, iż dorosłe osoby, które czują się na tyle silne, aby móc być niemoralne, jeśli są racjonalne, muszą przestrzegać zasad moralności. Jeśli znajdziemy się w sytuacji, w której nic nam nie grozi (np. staniemy się monarchą absolutnym w małym państwie, które czci nas jak Boga), bycie moralnym może nie być w zgodzie z egoistycznym interesem własnym. Jednak

najbezpieczniejszym i najlepszym sposobem wychowania naszego dziecka jest zaszczerpienie mu – o ile się potrafi – zespołu dobrych zasad moralnych oraz towarzyszących im uczuć – uczuć na tyle silnych, by zapewniły przestrzeganie tych zasad w zwykłych okolicznościach³¹³.

Ponieważ bycie niemoralnym egoistą wymaga zdolności, które dane są jedynie osobom wyjątkowo utalentowanym, a nie możemy mieć pewności, że wychowywane przez nas dziecko takie cechy będzie posiadało, najlepiej wychowywać je w duchu moralności³¹⁴. Z tego powodu wpojenie mu zasady: „Zawsze rób rzeczy niemoralne jeśli uważasz, że ci się uda” nie leży w jego interesie. Hare uważał również, iż ludzie, którzy przestrzegają zasad moralności, zazwyczaj są szczęśliwsi³¹⁵.

§ 4.14 Nadrzędność

Amoralizm, tak jak rozumiał go Hare, nie jest spreczny z prawami internalizmu motywacyjnego PIM1– PIM4³¹⁶. Są one bowiem regułami wewnętrznymi moralnej gry językowej i nie może ich podważyć osoba, która w ogóle nie używa (lub używa w niewłaściwy sposób) języka moralnego. Co więcej, nawet jeśli przyjęlibyśmy definicję amoralisty podaną przez D. Brinka, to amoralizm byłby spreczny

³¹⁰ Zob. tamże, s. 237.

³¹¹ Zob. tamże, s. 241.

³¹² Zob. tamże, s. 254.

³¹³ Tamże, s. 246.

³¹⁴ Tamże, s. 243.

³¹⁵ Zob. tamże, s. 252.

³¹⁶ Zob. § 4.13.

nie tyle z zasadami internalizmu Hare’a, ile z blisko z nim spokrewnionym prawem nadrzędności. Można bowiem twierdzić, że amoralista to osoba, która przedkłada inne motywy nad bycie moralnym i z tego względu nie akceptuje szczerze żadnych zasad etycznych.

Zasadę nadrzędności sformułować można w następujący sposób:

ZN Jeśli ktoś szczerze używa języka moralnego, to wszystkie zasady, które są w konflikcie z prawami moralnymi (np. zasady estetyczne, egoistyczny interes własny, wybory ze względów politycznych czy ekonomicznych), musi traktować jako podrzędne wobec zasad moralnych i w razie konfliktu je uchylić. „Traktowanie zasady jako nadrzędnej sprowadza się zatem do tego, że pozwalamy, by w razie konfliktu zawsze uchylała ona inne zasady, a tym samym, by uchylała wszelkie inne nakazy, także te, które nie poddają się uniwersalizacji (np. zwykłe pragnienia)”³¹⁷.

Prawo nadrzędności reguluje relacje pomiędzy moralnością a innymi praktykami oceniania: interesem własnym, estetyką, ekonomią, polityką itd. Głosi ono, iż reguły moralne muszą być nadrzędne względem innych reguł. Różnica pomiędzy internalizmem a nadrzędnością polega na tym, że jeśli internalizm jest wewnętrzną zasadą regulującą relacje pomiędzy danymi pojęciami moralnej gry językowej, to nadrzędność jest zasadą, której przedmiotem jest z jednej strony gra językowa moralności, z drugiej zaś strony inne praktyki oceniania. Według Hare’a nadrzędność nie jest jednak zasadą, która nie wchodzi w skład moralnej gry językowej i leży poza jej granicami. Już w *The Language of Morals* zauważył on, iż definiowanie sądów moralnych jedynie jako uniwersalnych preskrypcji nie jest adekwatne, gdyż zbiór taki jest zbyt szeroki, zawiera bowiem wypowiedzi, które nie jawią się jako moralne³¹⁸. Ponieważ uniwersalność i preskryptywność cechuje każdy dyskurs oceniający³¹⁹, wprowadził on zatem do definicji moralności *differentia specifica*. Jest nią właśnie nadrzędność³²⁰.

Sądy moralne = sądy oceniające (rodzaj) + nadrzędność (*differentia specifica*)

³¹⁷ Tamże, s. 75.

³¹⁸ Taki zarzut przeciwko Hare’a formalnej definicji sądu moralnego sformułował Mackie. Zob. J. L. Mackie, *Ethics...*, dz. cyt., s. 86.

³¹⁹ Krytykę tezy, iż oceny estetyczne podpadają pod prawo uniwersalizacji, znaleźć można w: D. A. Lloyd Thomas, *Aesthetics and Hare's Analysis of 'Good'*, „The Philosophical Quarterly” (56) 1964, s. 261-265.

³²⁰ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 38, 72.

Hare jednak nie był w tej kwestii stanowczy. Jego zdaniem,

dla wywodu nie będzie miało znaczenia, czy powiemy, że zastosowane w zasadzie sformułowanie „powinno się” oznacza to samo co wtedy, gdy zasadę potraktujemy jako nadrzędną, czy nie, tylko to, że ktoś, kto akceptuje zasadę, może, nie zmieniając znaczenia, traktować ją jako nadrzędną lub nie; nie będzie też miało znaczenia, czy powiemy, że jeśli traktuje ją jako nadrzędną, to zwrotu powinno się używać w inny sposób³²¹.

W *Mysleniu moralnym* rozszerzył on zakres sądów, które należy nazwać moralnymi. Nie są to jedynie uniwersalne zalecenia o charakterze nadrzędności, ale również te zasady, które choć nie są nadrzędne, traktowane są jako moralne, ponieważ jako swojego uzasadnienia wymagają namysłu krytycznego oraz zostały z niego wyprowadzone³²². Każdy sąd moralny, jeśli nawet nie jest sądem krytycznym o charakterze nadrzędności, musi czerpać z takiej zasady uzasadnienie i od niej pochodzić³²³.

Wprowadzając nadrzędność, Hare w jeszcze jednym punkcie odciął się od tradycji emotywistycznej. W tym bowiem kontekście dosłownie odczytana teza Ludwiga Wittgensteina, iż „etyka i estetyka to jedno”³²⁴, jest fałszywa. W aspekcie formalnym również dzięki nadrzędności język moralny nie jest tożsamy z innymi dyskursami normatywnymi, różni się bowiem nie tylko dziedziną stosowania, ale także własnościami formalnymi.

Z prawem nadrzędności związane są jednak pewne problemy. Choć zgoda na to, że oceny etyczne powinny uchylać oceny estetyczne czy ekonomiczne, jest raczej powszechna, wydaje się, iż analiza języka moralności nie dostarcza żadnej prawdy pojęciowej, która zdawałaby sprawę z relacji nadrzędności – uprzywilejowania dyskursu moralnego nad innymi praktykami oceniania. To, czy nadrzędność jest prawdą analityczną, aprioryczną, psychologiczną, czy socjologiczną, nie jest do końca jasne³²⁵. Pomimo iż Hare traktował

³²¹ Tamże, s. 76.

³²² Szczegółowe omówienie Hare’a ewolucji definicji moralności, wraz z krytyką pojęcia nadrzędności, znaleźć można w: W. Frankena, *Hare on Moral Weakness and the Definition of Morality*, „Ethics” (4) 1988, s. 779-792.

³²³ Zob. tamże, s. 80-81.

³²⁴ L. Wittgenstein, *Tractatus...*, dz. cyt., s. 80.

³²⁵ Pogląd, iż nadrzędność jest cechą analityczną, dla wielu wydaje się być mało atrakcyjny. Zob. T. Nagel, *The Foundation...*, dz. cyt., s. 112. Również J. Mackie atakuje Hare’ego, pisząc: „Mogłoby być tezą logiczną, iż terminy moralne mają takie znaczenie, że użyte w sądach czynią je uniwersalizowalnymi na drugi

ją jako cechę formalno-logiczną moralnej gry językowej, na poparcie tej tezy nie podał żadnego argumentu. Czy nie jest raczej tak, że postulat nadrzędności wynika z roli, jaką w obecnej kulturze ludzie faktycznie nadają zasadom moralnym – roli, do jakiej pretenduje etyka? W takim wypadku nadrzędność nie posiadałaby charakteru analitycznego – byłaby jedynie postulatem syntetycznym. Aby obronić formalność nadrzędności, można obrać inną drogę argumentacji: jeżeli nakazy etyczne są racjonalną, uniwersalną odpowiedzią na najbardziej podstawowe i najważniejsze dla nas pytania, gdybyśmy nie działali według najlepszych wskazań udzielonych w języku, w którym są one zadane, postępowałibyśmy nieracjonalnie. W ten sposób opisana moralność, jako zbiór nakazów roztropności, praw racjonalności praktycznej, implikowałaby naszą konieczność działania według tych zasad, jeśli chcielibyśmy być osobami racjonalnymi. Jednak powyższa odpowiedź, sprowadzająca bycie moralnym do bycia roztropnym, nie bierze pod uwagę Hare’a rozwiązania problemu amoralizmu. Jeśli bowiem amoralizm w interpretacji Hare’a nie jest nieracjonalny i wewnętrznie sprzeczny (można nie chcieć używać języka moralnego i być racjonalnym), nie wydaje się, aby mógł on utrzymać powyższe stanowisko. Co więcej, jeśli fenomen istnienia amoralistów o cechach, które przypisuje im D. Brink, jest prawdziwy, stanowi on dowód przeciwko formalnemu charakterowi nadrzędności sądów moralnych. Nadrzędność i amoralizm Brinka są bowiem niezgodne. Nie może istnieć osoba, która szczerze używa nadrzędnego języka moralnego oraz nie przestrzega wszystkich jego reguł (w tym prawa internalizmu motywacyjnego czy uniwersalizowalności); nie może nie popadać w sprzeczność osoba, która w swoich wyborach kieruje się, np. wąsko pojętym interesem własnym czy zmysłem estetycznym, zarazem wygłaszając szczerze nadrzędne sądy moralne. Z tego również względu Hare – ponieważ przyjmował tezę o formalności nadrzędności – nie zaakceptował możliwości istnienia amoralisty w interpretacji Brinka.

i na pierwszy sposób; ale to, iż maksymy, które przejdą ten test, muszą nas motywować do działania, byłoby materialną, praktyczną zasadą”; J. L. Mackie, *Ethics...*, dz. cyt., s. 92.

§ 5. Uniwersalny preskrytywizm

W niniejszym rozdziale przyjrzymy się kolejnym prawom pojęciowym odnoszącym się do moralnej gry językowej, jednak w przeciwieństwie do wcześniejszych analiz przedmiotem naszej uwagi będą przede wszystkim te cechy języka moralnego, które wspólne są ocenom i opisom. W związku z tym zajmiemy się zagadnieniem: uzasadnienia, uniwersalności oraz opiszemy stosunek Hare'a do kwestii prawdziwości i obiektywności sądów etycznych.

§ 5.1 Uzasadnienie

Paul Taylor¹ atakował pogląd głoszący, że oceny są preskrypcjami. Jego zdaniem, czynność zalecania jest różna od czynności ocenia, preskrypcja bowiem nie wymaga podania uzasadnienia, ocena zaś musi być podparta racjami. Posługiwał się on jednak węższym znaczeniem zalecania niż Hare. Hare bowiem zgodziłby się, iż sądy moralne wymagają uzasadnienia, przeczyłby jednak, iż z tego powodu nie mogą być one zaleceniami. Dzieje się tak, ponieważ nie są one, jak uważał P. Taylor, preskrypcjami typu A i A', lecz B, C lub D². Między innymi ta własność odróżnia sądy moralne od imperatywów: imperatyw A, w przeciwieństwie do zasady B, nie wymaga podania racji uzasadniającej. Choć w praktyce często zakładamy, że wypowiedziany nakaz może mieć racjonalne podstawy, jednak logika imperatywów A i A' nie wymaga konieczności ich istnienia. Rozkaz „Zapłać mu pieniądze!”, choć ma swoją przyczynę, nie musi

opierać się na racji. Zdaniem Hare'a, jeśli uważamy, że nadawca owego rozkazu na mocy pojęciowej konieczności musi mieć rację na jego poparcie, w gruncie rzeczy traktujemy jego wypowiedź A jako C³. Racją uzasadniającą dla zaleceń typu C i D jest zasada typu B. „Podanie preskrypcji C i zaprzeczenie, że istnieje jakaś zasada, na której ona się opiera, byłoby logicznie nieuprawnione”⁴. Użycie słowa „powinno się” bez wzięcia pod uwagę powyższych implikacji byłoby na tyle ekscentryczne, iż nie rozpoznałibyśmy tego słowa jako powinności, która jest różna od rozkazu⁵. Zasada przedstawiona powyżej dotyczy zarówno słowa „powinno się” użytego w zaleceniach typu C i D o charakterze moralnym, jak i pozamoralnym: różnica pomiędzy „Powinieneś podać podwójną dawkę trucizny” (wypowiedź skierowana do truciela) oraz „Powinieneś powiedzieć prawdę” polega jedynie na tym, że pierwsza z nich odnosi się do reguły określającej, jak być dobrym truciela, druga natomiast najprawdopodobniej do zasady o charakterze moralnym. Obie jednak wymagają istnienia reguły typu B. Z samego faktu, iż zalecenia C i D wymagają istnienia reguły typu B jako racji uzasadniającej, nie wynika, że mają one charakter moralny. To, czy wypowiedź C lub D jest moralna, zależy od założonego przez nadawcę charakteru normy B, która leży u jej podstaw. Nawet w przypadku wypowiedzi „On jest dobrym człowiekiem” nie możemy być pewni jej moralnego charakteru, jeśli nie znamy kontekstu, w jakim została ona wypowiedziana. Racją takiego sądu może być bowiem np. norma roztropnościowa czy religijna⁶. Zatem to moralne zasady typu B wyznaczają dziedzinę wypowiedzi moralnych⁷. Aby uzasadnić zasadę typu B, nie wystarczy jednak tylko znajomość wszystkich faktów, które są z nią związane, ani znajomość żadnej innej zasady moralnej, która byłaby od niej bardziej podstawowa. Proces jej uzasadnienia jest skomplikowany i wymaga odwołania się do pojęcia racjonalnego wyboru, które wraz z cechami języka moralnego funduje Hare'a utilityzm reguł – preferencji.

³ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 156-157.

⁴ Tamże, s. 157.

⁵ Zob. tamże.

⁶ Por. tamże, s. 155-161.

⁷ Zob. tamże, s. 160-162.

¹ Zob. P. W. Taylor, *Prescribing and Evaluating*, „Mind” (71) 1962, s. 213-230.

² Zob. tabela 3 na s. 81.

§ 5.2 Uniwersalność

Według Hare'a, reguły etyczne różnią się od innych zasad (np. zasad skutecznego trucicielstwa czy etykiety) typem uniwersalności oraz nadrzędnością – są one „zasadami postępowania ludzi jako ludzi”, nie zaś regułami wybranych grup społecznych⁸. Za kolejną po nadrzędności różnicę pomiędzy zasadami moralnymi typu B⁹ a zasadami pozamoralnymi odpowiada między innymi uniwersalność. Każda preskrypcja, która jest moralna, jest bądź uniwersalna, bądź uniwersalizowalna (podatna na uniwersalizację)¹⁰. Jednak nie jest to ta sama cecha, która odróżnia imperatyw A od A'. Aby przedstawić uniwersalność zasad typu B, należy przeciwstawić ją ogólności, odróżnić uniwersalność imperatywów od ocen (uniwersalność niepełną i pełną) oraz podzielić na formalną i materialną.

Uniwersalność zasad etycznych różni się, o czym niektórzy z komentatorów Hare'a zapominają¹¹, od ogólności. Aby to wyjaśnić, Hare przedstawił ich opozycje: opozycją ogólności jest szczegółowość, uniwersalności zaś jednostkowość¹². Ogólność jest parametrem opisującym ilość oraz szczegółowość predykatów użytych w danym sądzie, tak więc w przeciwieństwie do uniwersalności, jej cechą jest stopniowalność¹³. Zasada uniwersalna może być na tyle złożona i szczegółowa, że człowiek nie będzie w stanie jej zapamiętać¹⁴. Jest ona jednak uniwersalna, jeśli nie zawiera kwantyfikatorów jednostkowych, a wszystkie występujące w niej predykaty związane są kwantyfikatorem ogólnym¹⁵. **Uniwersalny sąd moralny** to taki akt mowy, który jest zaleceniem bez stałych indywidualnych (np. nazw osób, miejsc, czasu), wyrażonym jedynie za pomocą predykatów odnoszących się do własności uniwersalnych. **Własność** jest

⁸ Por. tamże, s. 162.

⁹ Różne rodzaje preskrypcji przedstawione zostały w § 4.1 (zob. tabela 3 na s. 81).

¹⁰ Tenże, *Myślenie...*, dz. cyt., s. 87. Uniwersalizowalność jest cechą zaleceń polegającą na możliwości uczynienia z nich w pełni uniwersalnych sądów moralnych.

¹¹ Pomijanie tego rozróżnienia przyczyniło się do wysuwania przeciwko uniwersalnemu preskrytywizmowi nieuzasadnionych zarzutów.

¹² Zob. tamże, s. 57; tenże, *Freedom...*, dz. cyt., 39-40.

¹³ Zob. tamże, s. 40.

¹⁴ Por. tamże, s. 39.

¹⁵ Zob. tamże, s. 176.

uniwersalna, jeśli do jej charakterystyki nie potrzebujemy odwołać się do żadnego konkretnego indywiduum¹⁶. Według Hare'a, gdy pojawi się jakikolwiek spór wokół tego, jakie własności są uniwersalne, należy zastosować następujący test: jeżeli dwie osoby, biorąc pod uwagę tę samą własność sytuacji, mogą niesprzecznie na jej podstawie wydać dwie odrębne oceny moralne, znaczy to, że własność owa nie jest uniwersalna¹⁷. Własności uniwersalne bywają nazywane również własnościami generalnymi lub wspólnymi¹⁸.

Rodzaj uniwersalności występujący w sądach moralnych i w imperatywach jest różny. Z tego względu Hare uniwersalność podzielił na pełną i niepełną. Zasady moralne typu B, C i D są **uniwersalne pełnie**, imperatywy A' przybierają mogą natomiast jedynie formę **uniwersalnych niepełnie**¹⁹. Jeśli w jakimś pomieszczeniu zauważymy zakaz A' „Nie palić!”, nie pomyślimy, że nie wolno palić również w innych miejscach, ponieważ są one dokładnie takie same jak pomieszczenie, w którym ten zakaz się znajduje. Według Hare'a, tzw. imperatywy uniwersalne A' („Nie palić!”, „Nie brudzić!”, „Nie hałasować!”) nie są uniwersalne we właściwym sensie tego słowa, są jedynie uniwersalne niepełnie. Można je nazwać uniwersalnymi, gdyż każdej osobie nakazują wykonać dany czyn, jednak ich zakres nie rozciąga się na sytuacje, które są istotnie podobne do pierwowzoru. Sąd moralny natomiast pociąga za sobą dalsze konsekwencje. Wypowiedź typu A': „Nie palić w tym pomieszczeniu!” oraz wypowiedź typu B użyta w znaczeniu ewaluatywnym: „Nie powinno się palić w tym pomieszczeniu” implikują: „Nie pal teraz w tym pomieszczeniu!”, jednakże jedynie zalecenie B pociąga za sobą: „Nikt nie powinien palić w żadnym pomieszczeniu dokładnie takim jak to”. („Nie pal w tym pomieszczeniu ani w dokładnie takim jak to!”²⁰). Taka pojęciowa zależność występuje jedynie w sądach uniwersalnych pełnych, imperatywy A oraz A' są jej pozbawione.

¹⁶ Zob. tenże, *Sorting...*, dz. cyt., s. 23.

¹⁷ Zob. tenże, *Myślenie...*, dz. cyt., s. 146.

¹⁸ Jacek Jadacki w następujący sposób definiuje własności generalne: „(x)(y)(P) [P jest własnością wspólną x-a i y-a \equiv (Px & Py)]”; J. J. Jadacki, *Spór o granice istnienia*, Wydział Filozofii i Socjologii UW, Warszawa 1998, s. 31.

¹⁹ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 175-179.

²⁰ Por. tamże, s. 178.

Zdaniem Hare'a uniwersalność pełna jest formalną własnością języka moralnego²¹. Jest to kolejne fundamentalne prawo logiki pojęć moralnych: jeśli dany przedmiot *A* nazwę dobrym lub moralnie pożądanym, wynika z tego, że każdy inny przedmiot, który jest taki sam jak *A*, według mnie jest również dobry lub pożądanym. To, iż uniwersalność jest zasadą formalną, mającą charakter pojęciowej konieczności²², Hare uzasadniał, odwołując się do kompetencji językowej:

Przyjmując, że *A* i *B* są dwoma indywiduami, nie można powiedzieć w sposób logicznie niesprzeczny, że w pewnej sytuacji, określonej w terminach ogólnych bez odniesienia do indywiduów, *A* powinien działać w sposób określony także w terminach ogólnych, lecz *B* nie musi działać w podobny sposób w podobnej sytuacji. [...] Znaczący to, że jeśli mówię: „tak właśnie powinno się postąpić, lecz może zdarzyć się sytuacja pod względem jej cech pozamoralnych dokładnie taka sama, w której jednak jakaś osoba – dokładnie ta sama osoba, która powinna to w tej sytuacji uczynić – nie powinna tego robić” – zaprzeczam sam sobie²³.

Prawo to bywa często nazywane mianem superweniencji. **Superweniencję** predykatów etycznych na predykatów naturalnych zazwyczaj wyraża się tezą, iż jeśli dwa byty są nierozróżnialne w odniesieniu do ich predykatów naturalnych, to w odniesieniu do nich stosują się również te same predykaty moralne. Można wyróżnić dwa rodzaje moralnej superweniencji (Sup)²⁴:

Sup (I) Jest konieczne, że jeśli jakiś przedmiot jest Nat^{*25} i Mor^{26} , to każdy przedmiot, który jest Nat^* , jest Mor . Konieczność rozumiemy w tym miejscu jako konieczność pojęciową. Ktoś, kogo wypowiedzi moralne nie respektują tego prawa, jest osobą, która źle używa pojęć moralnych.

Sup (II) Każdy przedmiot, który jest Nat^* , z konieczności musi być Mor . W tym miejscu użyta konieczność nie jest pojęciowa, lecz metafizyczna. Możemy powiedzieć, nie przeciwstawiając się regułom języka, że np. Hitler był osobą dobrą lub złą, jednakże tylko jedna z tych ocen z konieczności jest prawdziwa.

²¹ Zob. tenże, *Myślenie...*, dz. cyt., s. 146-147.

²² Zob. tenże, *Freedom...*, dz. cyt., s. 37.

²³ Tenże, *Uniwersalny...*, s. 504.

²⁴ S. Blackburn, *Supervenience...*, dz. cyt., s. 60-61.

²⁵ Nat^* jest zupełną charakterystyką przedmiotu dokonaną za pomocą predykatów naturalnych.

²⁶ Mor jest predykatem moralnym.

Hare odrzucił ewentualność drugą²⁷. Bliższe mu było prawo Sup (I), które sformalizować można w następujący sposób²⁸:

$$\text{Sup (I)} \quad \Box((\exists x)(Morx \ \& \ Nat^*x \ \& \ (Nat^*x \ \text{Sup} \ Morx)) \ \Box \ (y) \ (Nat^*y \ \Box \ Mor y))$$

Zdaniem Dalii Draii²⁹, prawo superweniencji, którym posługiwał się Hare, głosi, iż „każde przypisanie predykatu moralnego Mor presuponuje istnienie reguły $(x) \ (Natx \ \Box \ Morx)$ ”³⁰. Jeśli osoba *O* opisze czyn *x* w jego aspekcie naturalnym jako Nat_1 oraz zaleci go na mocy Nat_1 (przypisze mu predykat moralny Mor), zgodnie z zasadą superweniencji jest ona logicznie zobowiązana do zalecenia każdego innego czynu *y*, który posiada własności Nat_1 ³¹. D. Draii nazwała zasadę tę prawem **superweniencji askryptywnej**, aby odróżnić ją od superweniencji ontologicznej (np. Sup (II)), która dotyczy metafizycznych własności przedmiotów, a nie procesu uzasadniania. Jednak według niej nie jest prawdą, że uzasadnienie z konieczności podlegać musi temu prawu – nie każde uzasadnienie oceny wymaga istnienia reguły łączącej konkretne predykaty naturalne z konkretnymi predykatami moralnymi. Prawo superweniencji askryptywnej nie jest więc formalne – nie jest prawdą pojęciową języka moralnego. Aby uzasadnić jakąś ocenę, „nie musimy odnosić się ani do ogólnych zasad, ani nawet zakładać, iż one istnieją”³². Nie jest tak, że jeśli osoba O_1 jest Mor z racji tego, że ma pewne cechy Nat , to z konieczności osoba O_2 , która również je posiada, także musi być Mor . Może bowiem zdarzyć się tak, że osoba O_2 dodatkowo posiada inne cechy (jest np. chciwa i nieczuła), które powodują, że nie przypiszemy jej predykatu Mor ³³. Argument ten zrekonstruować można w następujący sposób:

²⁷ Zob. R. M. Hare, *Supervenience...*, dz. cyt., s. 71. Uwagi natury formalnej do sposobu przedstawienia superweniencji przez Hare'ego w zmiankowym artykule znaleźć można w: A. J. Dale, *Hare on Supervenience*, „Mind” (376) 1985, s. 599-600.

²⁸ Por. S. Blackburn, *Supervenience...*, dz. cyt., s. 60.

²⁹ D. Draii, *Moral Supervenience and Moral Thinking*, „Disputatio” (8) 2000.

³⁰ Aby cytował pasował do wywodu, autor niniejszej pracy zmienił zmienne: *A* na Nat i *B* na Mor . Zob. tamże, s. 21.

³¹ Zob. tamże, s. 23.

³² Tamże, s. 26.

³³ Zob. tamże, s. 23. Identyczny argument wysunął wcześniej J. Dancy; J. Dancy, *On Moral Properties*, „Mind” (90) 1981, s. 376-378. Poddał on krytyce pojęcie

1. Osoba O_1 , oceniająca fakt Nat^* jako Mor_1 , czyni to na podstawie Nat_1
2. Z tego na mocy prawa superwencji askryptywnej $[(x) (Natx \sqsupset Morx)]$ wynika, że każdy fakt Nat^* jest Mor_1 .

Jednak okazuje się, że:

3. Dla osoby O_2 , pomimo iż $Nat_1 \subset Nat^*$, istnieje Nat_2 , takie że $Nat_2 \cup Nat_1 \subset Nat^*$, i owo Nat_2 stanowi o tym, że osoba O_2 Nat^* ocenia jako $\sim Mor_1$.
4. Po zastosowaniu superwencji askryptywnej do tezy (3) wynika z niej oraz z (2), że $Nat^* \sqsupset Mor_1 \wedge Nat^* \sqsupset \sim Mor_1$
5. Jako że (3) jest powszechnie znanym fenomenem, powyżej zaprezentowane prawo superwencji askryptywnej nie jest prawem formalnym, gdyż prowadzi do sprzeczności (4).

Jednak zasada uniwersalizowalności Hare'a nie do końca zgadza się z superwencją zaprezentowaną przez D. Draia oraz S. Blackburna (Sup (I)). Blackburn przedstawił ją bowiem jako konieczną relację pomiędzy zbiorem sądów dotyczących wszystkich naturalnych cech jakiejś sytuacji (Nat^*) a moralną oceną Mor , Draia jako relację uzasadniania istniejącą pomiędzy oceną Mor a danym podzbiorem $Nat \subset Nat^*$, dla Hare'a natomiast superwencja jest koniecznym stosunkiem pomiędzy istotnymi własnościami uniwersalnymi – zbiorem wybranych, relewantnych własności naturalnych ($RNat$), które są racją dla osoby O do oceny moralnej Mor ³⁴. Regułę Sup I nie tylko powinniśmy ograniczyć do konkretnych własności relewantnych $RNat$, ale również zrelatywizować ją do danego nadawcy O . Tylko bowiem wtedy mieć ona będzie charakter formalnego prawa pojęciowego³⁵. Uniwersalność dotyczy jedynie istotnych własności natu-

superwencji i uniwersalizowalności, próbując wykazać, iż bardziej odpowiednim terminem opisującym relację między własnościami naturalnymi a moralnymi jest *resultance*. Podstawową różnicą między *resultance* a superwencją jest to, iż termin pierwszy opisuje relację zachodzącą między własnością moralną Mor a niektórymi własnościami pozamoralnymi, decydującymi, iż dany czyn jest właśnie Mor , podczas gdy drugi między wszystkimi własnościami naturalnym Nat^* danego czynu a własnością moralną Mor . Jednak jeśli przyjmiemy takie rozróżnienie, Hare'owi bliższe byłoby pojęcie *resultance*. Zob. poniżej przedstawione prawo HSup I.

³⁴ Zob. R. M. Hare, *Freedom...*, dz. cyt., s. 33; tenże, *Myślenie...*, dz. cyt., s. 193-194; tenże, *Universalisability*, w: tenże, *Essays on the Moral...*, dz. cyt., s. 15.

³⁵ Błędna interpretację uniwersalnego preskrytywizmu, w której nie bierze się pod uwagę relatywizacji prawa superwencji do osoby O można, znaleźć

ralnych, gdyż, jak zauważył J. L. Mackie, byłaby ona trywialna lub bezużyteczna, gdybyśmy nie potrafili eliminować przy ocenianiu danych przypadków nieistotnych moralnie różnic, abstrahować od błahych dla myślenia moralnego własności³⁶. Zdaniem Hare'a, jeśli oceniam jakiś przedmiot czy zdarzenie, nie mogę bez złamania reguł języka powiedzieć, że: „ A jest dobre/piękne, ale B , które jest identyczne jakościowo z A , nie jest dobre/piękne”. Podobna zależność zachodzi również wtedy, gdy mówimy o takim podobieństwie A i B , które jest wyznaczone jedynie przez istotne dla danej oceny własności uniwersalne. Na przykład nie mogę bez złamania reguł języka powiedzieć, że A jest dobre, ale B , które ma wszystkie istotne cechy identyczne z A , nie jest dobre. Jeśli wypowiem sąd moralny typu C: „W tej sytuacji nie powinieneś pobrać organów od osoby O , ponieważ $RNat_1$ ” (uznajemy za relewantną cechę $RNat_1$, np. to, że pacjent O ma 5% szans na przeżycie po wypadku), na mocy pojęciowej konieczności muszę uznać, że zawsze, gdy jakiś pacjent ma 5% szans na przeżycie i nie zaistnieją inne relewantne moralnie cechy sytuacji, które powinniśmy wziąć pod rozwagę, nikt nie powinien pobierać od niego organów. Po uwzględnieniu powyższych zmian uniwersalność formalną, **superwencję w wersji Hare'a** (HSup I), można przedstawić następująco:

$$\text{HSup I} \quad \begin{aligned} & \square((\exists x)(Mor_{Ox} \& RNat_{Ox} \& (RNatx \text{ Sup } Morx)) \square \\ & \square (y) (RNat_{Oy} \square Mor_{Oy})) \end{aligned}$$

w: D. Locke, *The trivializability...*, dz. cyt., s. 25-44, oraz F. Jackson, *From Metaphysics to Ethics: A Defense of Conceptual Analysis*, Oxford, Clarendon Press 1998. Według F. Jacksona uniwersalny preskrytywizm jest tezą „o funkcjonalnej roli słuszności – słuszność jest własnością, którą możemy przepisać uniwersalnie”; tamże, s. 143. Jego zdaniem nakazy, zalecenia superwenują na własnościach naturalnych: „dwa światy dokładnie takie same pod względem własności deskryptywnych są dokładnie takie same pod względem tego, co jest zalecane, w jaki sposób i przez kogo”; tamże, s. 152. Jednak powyższa interpretacja czyni z autora *Myślenia moralnego* zwolennika przejścia od „jest” do „powinien”, z czym Hare by się nie zgodził. Zob. § 4.12. Jest ona nietrafna także z innego powodu: Hare nigdy nie darzył szacunkiem analiz metafizycznych (konieczności metafizycznej), a taki charakter ma odczytanie uniwersalnego preskrytywizmu przez F. Jacksona. D. Lock uważał natomiast, że Hare starał się udowodnić logiczny charakter prawa, które głosi, iż jeśli osoba O zaleca x , musi ona również uważać, że każdy na jej miejscu powinien zalecać x . Zob. D. Locke, *The trivializability...*, dz. cyt., s. 34.

³⁶ Zob. J. L. Mackie, *Ethics...*, dz. cyt., s. 83.

Krytyka D. Drai nie uderza w powyższą regułę, gdyż z powodu relatywizacji prawa HSup I do osoby O oraz istotnych cech $RNat$ jej argumentację należałoby zrekonstruować w odmienny sposób. Należy stwierdzić, że ocena Mor_1 dla osoby O_1 superweniuje na innych istotnych własnościach uniwersalnych $RNat_1$, niż ocena dla osoby O_2 , która oceniła przedmiot Nat^* jako Mor_2 ze względu na $RNat_2$ ($RNat_1 +$ chciwość $+ nieczułość$). Powoduje to, iż osoba O_1 i O_2 oceniają nie jeden przedmiot Nat^* , lecz dwa różne stany rzeczy $RNat_1$ i $RNat_2$, pomimo że $RNat_1 \cup RNat_2 \subset Nat^*$. W takiej sytuacji nie można jednak powiedzieć, że $Nat^* \sqsupset Mor_1 \wedge Nat^* \sqsupset \sim Mor_1$, lecz jedynie $RNat_{1(O1)} \sqsupset Mor_{1(O1)} \wedge RNat_{2(O2)} \sqsupset Mor_{2(O2)}$.

Może się jednak wydawać, że prawo superweniencji HSup (I) z powodu jego relatywizacji do osoby O prowadzi do etycznego subiektywizmu. W taki sposób krytykował Hare'a A. MacIntyre. Jego zdaniem, Hare zakłada, że

to, co nazywam dobrem i co uznaję za swoją powinność, zależy od *mojego* wyboru pewnych fundamentalnych ocen i [...] nie ma logicznych granic określających zakres ocen, spośród których mogę wybierać. Innymi słowy, preskrytywizm Hare'a jest ostatecznie odnowieniem poglądu, że moich wartościowań moralnych nie wspiera i nie może wspierać żaden większy autorytet od autorytetu moich wyborów. Zrozumieć pojęcia wartościujące oznacza zrozumieć, że nasze użycie tych pojęć samo w sobie nie wymusza na nas przyjęcia żadnego określonego zespołu poglądów moralnych. [...] Jest to powtórzenie poglądu Kanta o podmiocie moralnym jako prawodawcy, ale teoria ta czyni go w ten sposób arbitralnym suwerenem, który jest twórcą głoszonych przez siebie praw i który czyni je prawami, nadając im formę powszechnie obowiązujących nakazów³⁷.

Gdyby bowiem przyjąć rozważane prawo superweniencji HSup (I), każdy z podmiotów, wypowiadając sądy moralne, tworzyłby uniwersalne prawa, lecz nie byłoby jasne, dlaczego jego prawa miałyby obowiązywać osobę trzecią, skoro ona również potrafi wyprowadzić konkurencyjne zasady uniwersalne. Podobną obiektywizm wysunęła również P. Foot³⁸. Według niej, teoria Hare'a opiera się na dwóch problematycznych założeniach³⁹:

³⁷ A. MacIntyre, *Krótko...*, dz. cyt., s. 326-327.

³⁸ Analogiczne argumenty znaleźć można w: P. J. McGrath, *R. M. Hare: A prescriptive Theory of Ethics*, „Philosophical Studies” 14 (1965); J. C. Mackenzie, *Prescriptivism and Rational Behaviour*, „Philosophical Quarterly” (18) 1968; H. Veatch, *Good Reasons and Prescriptivism in Ethics: A Metaethical Incompatibility?*, „Ethics” 80 (1970); W. Lyons, *Is Hare's Prescriptivism Morally Neutral?*, „Ethics” (3) 1972, s. 259-261.

³⁹ Por. P. Foot, *Moral...*, dz. cyt., s. 197.

1. Można bez błędu logicznego oprzeć swoją ocenę danego przedmiotu Nat^* o własności $RNat$, których nikt inny nie potraktowałby jako racji.

2. Nawet jeśli większość ludzi uważa, iż jakieś własności $RNat$ są istotne i wystarczają do wyprowadzenia konkluzji moralnej Mor , może istnieć osoba O , która je odrzuci, gdyż dla niej nie będą stanowić żadnych racji.

Prawdziwość powyższych tez – argumentuje następnie Foot – jest bardzo problematyczna, bo nasza intuicja podpowiada nam, że pewne racje uzasadniające konkretne oceny są obiektywne. Na przykład kryteria użycia zwrotu „dobry koń” nie mogą być arbitralnie ustalone przez dowolny podmiot, są one bowiem określone społecznie i niezależne od woli pojedynczych osób⁴⁰. Ktoś, kto powiedziałby, że dobry koń to jedynie takie zwierzę, które jest szczerbate, nie rozumiałby znaczenia zwrotu „dobry koń”. Zatem aby uniknąć subiektywizmu czy relatywizmu, należałoby uznać, że tylko jeden zbiór istotnych dla danej oceny własności naturalnych $RNat \subset Nat^*$ jest naprawdę moralnie relewantny.

Zdaniem Hare'a, pomimo że pomiędzy kryteriami oceny (własnościami $RNat$) a oceną nie istnieje więź inferencyjna, wybory moralne nie muszą być arbitralne. Choć jest prawdą, że w znaczeniu słowa „dobry” nie ma niczego, co łączyłoby je raczej z jednym „świadectwem” niż z drugim⁴¹, to jeśli zależy nam na prawdziwości naszej wypowiedzi, świadectwa $RNat$ nie mogą być dobierane dowolnie⁴². Ich wyznaczenie podlega jednak odrębnemu procesowi poznawczemu niesprowadzalnemu do analizy znaczenia⁴³. Obiektywność sądów etycznych zagwarantowana jest bowiem nie istnieniem moralnych faktów, lecz procesem rozumowania, które jeśli zostanie w pełni przeprowadzone, doprowadzi do uznania niezmiennej, jednej oceny danej sytuacji⁴⁴. Ów proces Hare nazwał myśleniem krytycznym, a jego rezultat zasadami krytycznymi. Zostaną one przedstawione w części III, poświęconej jego utylitaryzmowi. W tym miejscu wystarczy powiedzieć, że myślenie krytyczne jest

⁴⁰ Zob. tenże, *Goodness...*, dz. cyt., s. 216.

⁴¹ Zob. tenże, *Moral...*, dz. cyt., s. 196.

⁴² Por. R. M. Hare, *Universalisability...*, dz. cyt., s. 20-21.

⁴³ Zob. tenże, *Myślenie...*, dz. cyt., s. 82-85, 97-98.

⁴⁴ Zob. tenże, *Sorting...*, dz. cyt., s. 134.

metodą wyznaczania obiektywnych i racjonalnych ocen – zasad krytycznych, którym można przypisać wartość prawdy, opartych z jednej strony na własnościach logicznych języka moralnego oraz istotnych cechach ocenianej sytuacji, z drugiej zaś na założeniach dotyczących racjonalnego namysłu. Wszystkie te wymogi ograniczają wolność wyboru naszej prawodawczej woli – kryteria wydawania moralnych zaleceń zostają zawężone. Z tego powodu, chcąc dowiedzieć się, czy dany czyn jest moralny i czy powinniśmy go wybrać, nie możemy stosować metody arbitralnej, co zarzucali Hare'owi A. MacIntyre i P. Foot. Aby spełnić wymogi preskryptywności i uniwersalności, musimy posłużyć się rozumem, który ograniczy naszą wolność wyboru, oraz wziąć pod uwagę zastane fakty.

Wolność, którą cieszymy się jako istoty tworzące dyskurs moralny, jest wolnością rozumowania, tzn. dokonywania racjonalnych ocen moralnych; zaś reguły tego rozumowania, określone przez pojęcia występujące w pytaniach, na które udzielam odpowiedzi, uświadamiają nam, że – postępując racjonalnie – w najważniejszych kwestiach moralnych będziemy korzystać z naszej wolności tylko w jeden sposób⁴⁵.

W świetle teorii Hare'a zasady krytyczne, tj. obiektywne prawa moralne, do których staramy się dojść, myśląc krytycznie, ostatecznie muszą podlegać następującej regule:

$$\text{HSup}^*\text{II} \quad \begin{array}{l} \Box((\exists x)(\text{Mor}_A x \ \& \ \text{RNat}_A x \ \& \ (\text{RNat}_x \ \text{Sup} \ \text{Mor}_x)) \ \Box \ (y) (\text{RNat}_A y \ \Box \ \text{Mor}_A y)) \\ \sim \Box(\exists x) (\text{RNat}_A x \ \& \ \sim \ \text{Mor}_A x) \end{array}$$

HSup*II zrelatywizowane jest nie, jak HSup I, do jakiejś fizycznej osoby *O*, lecz do jedynego idealnego nakazodawcy – archanioła *A*, który w doskonały sposób, zawsze metodycznie, sumiennie i sprawnie dokonuje namysłu moralnego, biorąc pod uwagę jedynie faktycznie istotne cechy *RNat* danej sytuacji etycznej. Z powodu analogii do idealnego obserwatora, archanioła można określić również mianem „idealnego nakazodawcy” i „idealnego wykonawcy nakazów”⁴⁶. Myślenie krytyczne bywa interpretowane w myśl teorii idealnego obserwatora z racji wielu podobieństw pomiędzy tymi metodami⁴⁷. Istnieją jednak zasadnicze różnice, które nie pozwalają ich ze sobą

⁴⁵ R. M. Hare, *Myślenie...*, dz. cyt., s. 18.

⁴⁶ Zob. tamże, s. 61.

⁴⁷ Zob. D. A. J. Richard, *Prescriptivism*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 118-121.

łączyć⁴⁸. Sposób myślenia idealnego nakazodawcy opisujemy w części III. Odwołując się do wzorcowego sposobu myślenia moralnego Hare zabezpieczył się przed subiektywizmem i relatywizmem oraz mógł postulować obiektywność i prawdziwość sądów moralnych. Wierzył on bowiem, że każda osoba, która o danej sytuacji *Nat** myślałaby tak jak archanioł, wydałaby tę samą ocenę moralną *Mor*⁴⁹. Ocenę tę należałoby nazwać prawdziwą.

§ 5.3 Prawda i obiektywność

Podstawowy zarzut skierowany pod adresem teorii nondeskryptywistycznych głosi, iż twierdzenie, że sądy etyczne nie mogą mieć wartości logicznych i co za tym idzie, niemożliwe jest przy ich użyciu wnioskowanie, prowadzi przedstawicieli nondeskryptywizmu do subiektywizmu, którego konsekwencją jest relatywizm⁵⁰. Zakłada się bowiem, że bez pojęcia prawdy niemożliwe jest postulowanie obiektywności sądów etycznych. W powyższym toku argumentacyjnym nonkognitywistów z konieczności traktuje się jako antyrealistów, antyrealizm zaś definiuje się jako pogląd, że niemożliwe jest używanie predykatu prawdy w etyce⁵¹. Twierdzi się, że jeśli dyskurs moralny ma posiadać wszystkie semantyczne cechy pozwalające na to, aby mówić o prawdziwości lub fałszywości sądów moralnych, musi istnieć realna sfera bytów, nazywana własnościami lub faktami. Zatem chcąc mówić o prawdzie, trzeba poprzeć redukcjonistyczny bądź antyredukcjonistyczny realizm etyczny.

Richard Hare nie zgadzał się jednak na powyższe argumenty, zdawał sobie bowiem sprawę z tego, iż zakładają one korespondencyjną teorię prawdy. Takie pojmowanie prawdy prowadzi w etyce do problemów, czego przykładem jest *teoria globalnego błędu* J. L. Mackiego. Mackie był kognitywistą i antyrealistą: według niego, zdania moral-

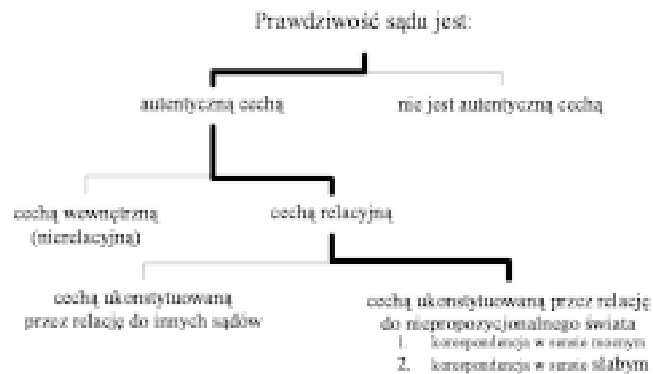
⁴⁸ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 61.

⁴⁹ Zob. tamże, s. 18, 64, 189, 278.

⁵⁰ „Przy takim ujęciu fenomenu moralności preskrytywista nie ma danych, które by umożliwiły wyróżnić określone czynniki wartościowości aktu. Ma do dyspozycji tylko uczuciowe stany. [...] W tych też końcowych sformułowaniach preskrytywizmu z całą wyrazistością dochodzi do głosu wspólna wszystkim trzem współczesnym stanowiskom podstawowa orientacja filozoficzna, jaką jest mniej lub więcej radykalny **etyczny relatywizm**”; T. Ślipko, *Zarys...*, dz. cyt., s. 157.

⁵¹ Por. C. Wright, *Truth in Ethics*, „Ratio” (8) 1995, s. 209.

ne odnoszą się do faktów, fakty moralne jednak nie istnieją, gdyż w przeciwnym razie musiałyby być bytami bardzo osobliwymi. Jego zdaniem, prawdziwość sądów etycznych sprowadza się do ich korespondencji z faktami, a w związku z tym żaden sąd moralny nie jest prawdziwy, ponieważ nie istnieją warunki, które mogłyby zostać spełnione przez relację prawdziwości. Korespondencyjna teoria prawdy doprowadziła go zatem do moralnego nihilizmu⁵². Dlaczego jednak mielibyśmy posługiwać się pojęciem prawdy, z którego wynikają tezy tak bardzo sprzeczne z naszymi intuicjami⁵³? Wydaje się, że nie musimy traktować reprezentacyjnego aspektu pojęcia prawdy w duchu **realizmu semantycznego**⁵⁴, prowadzi to bowiem do niepotrzebnych kłopotów. Zdaniem Hare'a, nondeskrytywizm uniknąć może zarzutu pozbycia się pojęcia prawdy z dziedziny etyki, jeśli uświadomimy sobie różnorodność akceptowanych współcześnie definicji prawdy. Owo bogactwo w schematyczny sposób za Crispinem Wrightem przedstawia rysunek 7⁵⁵.



Rysunek 7. Sposoby definiowania prawdy

⁵² Zob. J. L. Mackie, *Ethics...*, dz. cyt., s. 48-49.

⁵³ Por. C. Wright, *Truth in Ethics...*, dz. cyt., s. 211.

⁵⁴ **Realizm semantyczny** – znaczenie zdań sprowadza się do tego, co czyni owe zdania prawdziwymi, czyli do warunków prawdziwości. Owe warunki pomowane są przy tym jako zachodzące niezależnie od naszej wiedzy o nich i zdolności ich rozpoznania – przekraczają nasze świadectwo. Por. P. Gutowski, T. Szubka (red.), *Filozofia brytyjska...*, dz. cyt., s. 59.

⁵⁵ Por. C. Wright, *Prawda: przegląd tradycyjnego sporu*, „Kwartalnik Filozoficzny” (4) 2000, s. 158.

Filozofowie, którzy odrzucili nondeskrytywizm z powodu jego rzekomego problemu z prawdą i obiektywnością, posługiwali się tylko jedną teorią prawdy w specyficznej, mocnej interpretacji. Na powyższym rysunku przedstawiona jest ona jako gałąź oznaczona grubą kreską. Hare jednak nie uważał, iż prawda w etyce powinna być definiowana jako korespondencja sądu z faktami. Choć wykorzystywane w jego pracach pojęcie prawdy ewoluowało, stwierdzić można, że utrzymywało się w nurcie teorii deflacyjnych⁵⁶. Wyróżnić można dwa okresy w jego filozofii, w których odmiennie podchodził on do tej kwestii⁵⁷.

Przed napisaniem *Sorting Out Ethics* Hare o prawdzie mówił niewiele, w praktyce próbując zastąpić ją obiektywnością jako bezstronnością oraz racjonalnym uzasadnieniem⁵⁸. W związku z występującą wśród filozofów moralności tendencją do traktowania prawdy jako mocnej korespondencji oraz rozpowszechnionym założeniem, że jest ona cechą jedynie przekonań, w gruncie rzeczy zrezygnował z używania tego pojęcia. Po wprowadzeniu w *Myśleniu moralnym* dwóch płaszczyzn rozumowania moralnego⁵⁹, wyznaczył dziedzinę, w której można używać predykatu prawdy. Należą do niej zasady domniemane i płaszczyzna intuicyjna. Zasady domniemane, bliżej omówione w części III niniejszej książki, to nieskomplikowane, intuicyjne reguły moralne, których uczymy się w procesie socjalizacji. Podlegają one zmianom oraz są zależne od kontekstu społecznego i historycznego. Ponieważ pewne zasady domniemane posiadają powszechnie akceptowane, skonwencjonalizowane znaczenie deskrytywne, uważamy je za prawdziwe, jeśli są zgodne z powszechnie akceptowanymi warunkami ich stwierdzal-

⁵⁶ Wśród **deflacyjnych teorii prawdy** wymienia się: redundancyjną teorię prawdy Ramsaya, dyskwoatacyjną teorię prawdy Quine'a, performatywną teorię prawdy Strawsona, minimalną teorię prawdy Horwicha. Koncepcje te łączą przekonanie, że zwroty typu „to jest prawda” nie pełnią funkcji opisowych, lecz funkcje odmienne. Skrajni przedstawiciele deflacionizmu uważają, że za pomocą predykatu prawdy niczego się nie orzeka, ponieważ prawdziwość nie jest żadną autentyczną własnością czy relacją; zwolennicy umiarkowani natomiast, że jest to własność trywialna. Por. J. Woleński, *Epistemologia*, tom III: *Prawda i realizm*, Aureus, Kraków 2003, s. 85-91, oraz R. Ziemińska, *Antyrealizm i deflacionizm a pojęcie prawdy*, „Przegląd Filozoficzny” (2) 2000, s. 69-84.

⁵⁷ Zaznaczyć należy, że analiza Hare'a rozumienia prawdziwości sądów etycznych jest przedsięwzięciem trudnym, pisał on bowiem na ten temat bardzo rzadko, zebrane jego wypowiedzi łącznie zajmują tylko kilkanaście zdań.

⁵⁸ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 263-264.

⁵⁹ Zob. § 6.1.

ności – kryteriami ich użycia. Z tego powodu, choć zasady moralne są preskrypcjami, wydaje się, że mogą podlegać one parametrowi prawdy lub fałszu. Dzieje się tak za sprawą zawartego w nich znaczenia deskryptywnego⁶⁰. Aby podkreślić, że dana ocena jest zgodna z panującymi standardami, często nazywamy ją prawdziwą. To, czy zasady moralne nazywane są prawdziwymi, zależy od tego, czy spełniają one akceptowane warunki prawdziwości. Hare nie wyjaśnił bliżej, co przez nie rozumie, choć traktował je antyrealistycznie, tj. raczej jako warunki poprawnej stwierdzalności czy akceptowalności niż realistycznie pojmowane warunki prawdziwości przekraczające ludzkie świadectwo. Pojęciu prawdy na płaszczyźnie intuicyjnej nadał więc charakter epistemiczny⁶¹. Tak rozumiana prawdziwość jest zrelatywizowana do danej grupy lokalnej, która aprobejuje określone kryteria. W tym świetle obiektywność sądów moralnych jest rozumiana jako intersubiektywność, prawdziwość zaś sprowadza się do akceptowanego społecznie uzasadnienia. Z tego powodu prawdziwość intuicyjnych sądów moralnych nie jest własnością absolutną, lecz relatywną⁶².

Inaczej wygląda kwestia prawdziwości sądów, które wydałby idealny nakazodawca – archanioł. Do zasad myślenia krytycznego stosował on **performatywną teorię prawdy** Petera F. Strawsona⁶³. Położył nacisk na normatywny aspekt prawdy, negując zarazem jej charakter opisowy oraz to, że jest ona własnością autentyczną⁶⁴. Zgodnie z tą

⁶⁰ Zob. tenże, *The Language...*, dz. cyt., s. 28; tenże, *Freedom...*, dz. cyt., s. 28; tenże, *Sorting...*, dz. cyt., s. 138-140.

⁶¹ „Teoria prawdy jest epistemiczna, gdy sprowadza prawdziwość do uważania za prawdę”; J. Woleński, *Epistemologia...*, dz. cyt., s. 110. „Zwolennicy tych koncepcji utrzymują, że prawdziwość nie polega na jakiejś relacji do mniej lub bardziej «transcendentnego» stanu rzeczy czy faktu, lecz sprowadza się do tego, czy dane przekonanie lub twierdzenie jest *uzasadnione*, posiada *odpowiednie gwarancje*, jest *racjonalnie ugruntowane* itp. Tradycyjnym przykładem tych koncepcji prawdy są teorie koherencyjne, przypisywane takim autorom, jak F. H. Bradley czy B. Blanshard. Do epistemicznych można zaliczyć również pragmatyczne ujęcia prawdy, proponowane przez C. S. Peirce’a. W. Jamesa i J. Deweya. W ostatnich latach epistemicznej koncepcji prawdy bronił przez jakiś czas Hilary Putnam”; T. Szubka, *Realizm i minimalizm w teorii prawdy. Stanowisko W. P. Alstona*, „Kwartalnik Filozoficzny” (4) 1999, s. 116.

⁶² Zob. R. M. Hare, *Sorting...*, dz. cyt., s. 140.

⁶³ Zob. P. F. Strawson, *Prawda*, w: J. Pelc (wybrał i przełożył), *Logika i język. Studia z semiotyki logicznej*, PWN, Warszawa 1967, s. 353-375.

⁶⁴ Zob. R. M. Hare, *The Language...*, dz. cyt., s. 28; tenże, *Myślenie...*, dz. cyt., s. 262-264; tenże, *Some Confusions about Subjectivity*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 26-27.

koncepcją wyrażenie prawdziwościowe należy do klasy wyrażen nazywanych przez Austina performatywnymi. Gdy mówimy „jest prawdą”, nie opisujemy żadnej rzeczywistości pozajęzykowej, lecz dokonujemy działania polegającego na potwierdzeniu tego, co ktoś wcześniej stwierdził. Predykat „prawda” pełni więc niesemantyczną rolę potwierdzania. Mówiąc: „sąd x jest prawdziwy”, po prostu udzielamy mu poparcia⁶⁵. W ten sposób stwierdzanie prawdziwości staje się normatywną czynnością mowy: rekomendując jego przyjęcie, zalecamy go jako coś, co spełnia pewien standard. Na płaszczyźnie intuicyjnej czynimy to, gdy sąd moralny zgadza się z przyjmowanymi w danej grupie warunkami jego prawdziwości, natomiast na płaszczyźnie krytycznej, ponieważ jest on wynikiem zastosowania myślenia krytycznego, które jest metodą odkrywania najbardziej racjonalnej i obiektywnej oceny. Performatywna teoria prawdy jest jednak uwikłana w poważny problem przedstawiony przez G. Fregego, P. T. Geacha i J. Searle’a: jej zwolennicy mają duże trudności z wyjaśnieniem kontekstów złożonych czy podstawowych relacji inferencyjnych, takich jak *modus ponens*⁶⁶.

W drugim okresie, trwającym od lat dziewięćdziesiątych, podejście Hare’a do problemu prawdziwości sądów krytycznych nieznacznie się zmieniło. Zachował on natomiast koncepcję prawdziwości sądów domniemanych. Od tej chwili z pełną otwartością głosił, że „nondeskrytywista taki jak ja z łatwością może przystać, że sądy moralne mogą być prawdziwe lub fałszywe”⁶⁷. Pisząc w *Sorting Out Ethics* o prawdzie⁶⁸, powołał się na teorię przedstawioną przez Crispina Wrighta w *Truth and Objectivity*⁶⁹ – **minimalną koncepcję prawdy**.

Według C. Wrighta, nie powinno się faworyzować monistycznego ujęcia prawdziwości, lecz rozwijać **pluralistyczną koncepcję prawdy** (pluralizm alethyczny). Dotychczasowe teorie zakładały, że jest jedno właściwe jej pojęcie (monizm alethyczny), tymczasem może być wiele rodzajów prawdy dla różnych dyskursów. Klasyczna teoria korespondencyjna nie pasuje np. do twierdzeń matematyki (tu odpowiednia jest koncepcja koherencyjna) czy do dyskursu etycznego lub estetycznego (tutaj najwłaściwsza jest koncepcja epistemiczna). Po-

⁶⁵ Zob. tamże.

⁶⁶ Zob. § 4.7.

⁶⁷ Tenże, *Objective...*, dz. cyt., s. 17.

⁶⁸ Zob. tenże, *Sorting...*, dz. cyt., s. 58.

⁶⁹ C. Wright, *Truth and Objectivity*, Harvard University Press, Cambridge, Massachusetts 1992.

nieważ prawda jest pojęciem rodzinnym, nie jest możliwa jej analiza w kategoriach warunków koniecznych i wystarczających, która objęłaby jednocześnie różnorodność dyskursów⁷⁰.

Zdaniem Wrighta, być predykatem prawdy to spełniać zbiór pewnych ogólnych, intuicyjnych prawd apriorycznych, tzw. truizmów⁷¹. W *Objective Prescriptions* Hare, odwołując się do tego poglądu, opisał trzy cechy, które według niego niesie ze sobą pojęcie prawdy: pierwszą jest własność syntaktyczna, którą naświetlił A. Tarski, wyrażona w schemacie dyskwotacyjnym: „*p* jest prawdziwe $\square p$ ”; druga to funkcja udzielania poparcia, o której w 1949 roku pisał P. Strawson⁷²; trzecią z nich stanowi natomiast twierdzenie, iż „jest prawdziwe” jest ekwiwalentem „spełnia warunki prawdziwości, jakiegokolwiek by one nie były”⁷³. Pierwsze dwie własności zgodne są z tezami podejścia deflacyjnego⁷⁴. Natomiast trzecią cechą powinniśmy traktować jako stwierdzenie wykraczające poza deflacionizm, dzięki niemu bowiem prawda nie jest jedynie własnością trywialną, lecz niesie ze sobą element deskryptywny. Zdaniem Hare’a, owe trzy cechy są truizmami, które teoria prawdy sądów etycznych musi brać pod uwagę, przy czym nie stanowią one problemu dla uniwersalnego preskrytywizmu⁷⁵.

Richard Hare, powołując się na teorię prawdy Wrighta, nie określił, w jaki sposób miałyby ona współgrać z jego systemem. Także rozróżnienie dwu sposobów podejścia do prawdy w zależności od dwu płaszczyzn myślenia moralnego nie zostało przez Hare’a przedstawione *explicite*. Wydaje się jednak, że przedstawiona w tym miejscu interpretacja trafnie oddaje jego poglądy i intuicje. Nawet jeśli sam Hare nie podzielał do końca idei Wrighta, co trudno jest jednoznacznie stwierdzić, jego koncepcja prawdy może zostać spójnie złączona z uniwersalnym preskrytywizmem, i to z dużym dla niej pożytkiem. Jeśli powyższa interpretacja jest uzasadniona, mogłaby ona pomóc w oddaleniu wielu zarzutów skierowanych w stronę jego nondeskrytywizmu. Choć Hare nie sformułował jasno, które elementy teorii prawdy C. Wrighta akceptuje, a które odrzuca, wydaje się, iż

mógłby przyjąć ją w całości przy zastrzeżeniu, że potraktowalibyśmy ją jako teorię dotyczącą sądów myślenia krytycznego. Można zaproponować takie podejście, jeśli myślenie krytyczne przeprowadzone przez archaniola zinterpretujemy jako metodę dojścia do **superstwierdzalności**⁷⁶. Według Wrighta bowiem zdanie prawdziwe w sensie minimalnym to zdanie superstwierdzalne, czyli takie, które jest zasadnie stwierdzalne i takim pozostaje bez względu na możliwe rozszerzenia naszego zasobu informacji⁷⁷. „Moralnie prawdziwy byłby sąd, który jest moralnie uzasadniony i takim pozostaje bez względu na to, jak szczegółowe i doskonale podjęlibyśmy w tej kwestii dodatkowe rozważania”⁷⁸. Myślenie krytyczne przeprowadzone przez archaniola daje właśnie takie rezultaty⁷⁹. Dzieje się tak, ponieważ każdy idealnie racjonalny podmiot, postawiony w obliczu pozamoralnych faktów, zgodzić się musi na jedną ocenę. Obiektywność sądów etycznych jest bowiem zagwarantowana nie istnieniem moralnych faktów, lecz dzięki procesowi rozumowania, które jeśli zostanie przeprowadzone przy spełnieniu idealnych warunków, doprowadzi do uznania niezmiennej, jednej oceny danej sytuacji⁸⁰. Obiektywność, pojmowana przez Hare’a jako bezstronność⁸¹, zagwarantowana jest więc przez superstwierdzalność sądów moralnych. Z tej perspektywy również Hare mógłby przystać na następujące cechy, które Wright postuluje jako truizmy związane z pojęciem prawdy⁸²: przezroczystość prawdy (każda postawa wobec sądu jest postawą wobec jego prawdziwości); nieprzejrzystość prawdy (podmiot poznający może być tak usytuowany, że dana prawda znajduje się poza jego zasięgiem, tak że niektóre prawdy mogą nigdy nie zostać poznane); zachowanie zdolności do bycia prawdziwym w powiązaniach (zdolność do bycia prawdziwym jest zachowywana w rozmaitych logicznych operacjach⁸³); truizm korespondencyjny (prawdziwość sądu jest wyznaczona przez korespon-

⁷⁰ Por. tenże, *Prawda: przegląd...*, dz. cyt., s. 180.

⁷¹ Zob. tenże, *Truth...*, dz. cyt., s. 34.

⁷² P. F. Strawson, *Prawda...*, dz. cyt., s. 353-375.

⁷³ Por. R. M. Hare, *Objective...*, dz. cyt., s. 16; tenże, *Sorting...*, dz. cyt., s. 57-58.

⁷⁴ Por. C. Wright, *Prawda: przegląd...*, dz. cyt., s. 161-162.

⁷⁵ Zob. R. M. Hare, *Objective...*, dz. cyt., s. 17-18.

⁷⁶ Zob. tenże, *Truth in Ethics...*, dz. cyt., s. 218.

⁷⁷ Por. tamże, s. 217-218.

⁷⁸ Tamże, s. 218.

⁷⁹ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 64, 278.

⁸⁰ Zob. tenże, *Sorting...*, dz. cyt., s. 134.

⁸¹ Zob. tenże, *Myślenie...*, dz. cyt., s. 261.

⁸² Zob. C. Wright, *Prawda: przegląd...*, dz. cyt., s. 183-184.

⁸³ Przyjęcie tego truizmu mogłoby rozwiązywać problem wnioskowań z użyciem moralnych przesłanek oraz problem złożonych kontekstów gramatycznych (argument Fregego-Geacha).

dencję z rzeczywistością, trafne odzwierciedlanie tego, jak się rzeczy mają; kontrast prawdy i uzasadnienia (sąd może być prawdziwy, nie będąc uzasadnionym, podobnie jak może zostać uzasadniony, nie będąc prawdziwym); beczasowość prawdy (jeśli sąd jest kiedykolwiek prawdziwy, to jest zawsze prawdziwy); absolutność prawdy (prawdziwości nie można stopniować).

Zdaniem Wrighta, choć filozofowie moralności często uzależniają definicję realizmu i antyrealizmu moralnego od tego, czy sądy moralne posiadają wartości logiczne, to kwestie realności i prawdziwości nie muszą być od siebie zależne. Prawda minimalna jest neutralna wobec sporu realizmu z antyrealizmem i nie jest wyłączną własnością pierwszego z nich⁸⁴. Jeśli ktoś uważa, że prawda jest kwestią tego, co według nas jest akceptowalne w świetle pełnego oszacowania faktów pozamoralnych oraz pewnych niekognitywnych dyspozycji do odczuwania moralnego, nie musi zrezygnować z akceptacji jej wersji korespondencyjnej. Będąc nondeskrytywistami, możemy traktować prawdę jako pewną relację sądów do aspektów niepropozycjonalnego świata, jeśli nie będziemy jej pojmowali jako korespondencji w mocnym sensie⁸⁵. Korespondencyjną teorię prawdy podzielić można bowiem na kilka rodzajów, ze względu na to, co rozumie się przez korespondencję. Możemy mówić o korespondencji w sensie słabym (semantycznym) i w sensie mocnym⁸⁶. **Korespondencja w sensie słabym** w teorii prawdy sprowadza się do takich relacji semantycznych, jak referencja (oznaczanie) i spełnianie, które zachodzą pomiędzy nośnikiem prawdy a odpowiednimi przedmiotami. Jeśli semantyczne pojęcia referencji i spełniania zinterpretujemy minimalistycznie – jako odnoszące się do relacji zachodzących między pewnym językiem a dziedziną, o której wiadomo jedynie, że jest w jakimś sensie „zewnętrzna” wobec niego (ta dziedzina nie musi być obiektywna czy realna i bezpośrednio dostępna), to w konsekwencji korespondencja nie będzie miała żadnych zobowiązań epistemologicznych i metafizycznych, jakie postulują realiści moralni. Natomiast **korespondencja w sensie mocnym** musi spełniać warunki dodatkowe: wymaga określonej struktury sądu lub zdania; określo-

⁸⁴ Por. tamże, s. 180.

⁸⁵ Por. tamże.

⁸⁶ Por. T. Szubka, *Minimalistyczne a korespondencyjne teorie prawdy*, w: J. Hartman (red.), *Filozofia i logika. W stronę Jana Woleńskiego*, Aureus, Kraków 2000, s. 383 n.

nej struktury fragmentu rzeczywistości (stanu rzeczy, faktu lub sytuacji) oraz odpowiedności między obiema strukturami⁸⁷.

Choć Hare przeciwstawiał zalecenia stwierdzeniom czy opisom – kontrastował wypowiedzi o mocy illokucyjnej zalecenia z tymi o mocy stwierdzania – nie głosił, że sądy moralne nie mogą być wyrazem mentalnych stanów przekonaniowych czy wiedzy. Również te kognitywne cechy sądów moralnych, obok posiadania wartości logicznych, nie były przez niego odrzucane. Stany mentalne skorelowane z sądami moralnymi można nazwać przekonaniem, utrzymując, że są one innym ich rodzajem niż przekonania faktualne⁸⁸. Ponieważ każdy sąd moralny, zarówno domniemany jak i krytyczny, posiada znaczenie deskryptywne, warunki prawdziwości oraz niesie z sobą informację⁸⁹, przekonanie możemy pojmować jako stan umysłu skorelowany ze znaczeniem deskryptywnym. Różnica między sądami ewaluatywnymi a opisowymi polega bowiem jedynie na tym, że pierwsze posiadają coś więcej niż drugie. Sądy ewaluatywne składają się z:

- (1) elementu syntaktycznego, który determinuje (2) ich moc illokucyjną (to, że są sądami ewaluacyjnymi), co z kolei wymaga (3) posiadania przez nie warunków prawdziwości; mają one (4) konkretne warunki prawdziwości; ich moc illokucyjna wymaga, (5) aby były ocenami, co z kolei oznacza, że mogą posiadać ewaluacyjną moc illokucyjną nawet wtedy, gdy ich warunki prawdziwości ulegają zmianie⁹⁰.

Nondeskrytywiści nie muszą więc – co potwierdza Kyle S. Swan – definiować swojej teorii za pomocą tez semantyczno-gramatycznych, które wskazują na niepodatność na prawdę czy niewyrażanie przekonania⁹¹, lecz mogą odwoływać się jedynie do pragmatyki: mówić o czynnościach mowy, które są wykonywane podczas standardowego wypowiedzania sądów moralnych⁹². Z tego względu tacy nondeskrytywiści jak Hare mogą być kognitywistami. Deflacyjna czy minimalna⁹³ teoria prawdy dla nondeskrytywistów jest bardzo

⁸⁷ Por. tamże.

⁸⁸ Zob. R. M. Hare, *Ontology...*, dz. cyt., s. 96-97.

⁸⁹ Zob. tenże, *Objective...*, dz. cyt., s. 17-18.

⁹⁰ Tenże, *Sorting...*, dz. cyt., s. 59.

⁹¹ Takie uproszczenie znaleźć można w szeroko dyskutowanym artykule: F. Jackson, P. Pettit, *A Problem for Expressivism*, „Analysis” (4) 1998.

⁹² Zob. K. S. Swan, *Emotivism and Deflationary Truth*, „Pacific Philosophical Quarterly” (83) 2002, s. 280.

⁹³ Niekiedy pojęcia minimalnej i deflacyjnej teorii prawdy są stosowane zamiennie, niekiedy natomiast minimalną teorię prawdy Horwicha (P. Horwich, *Truth*,

atrakcyjna, ponieważ pozwala odeprzeć wiele stawianych im zarzutów. Przede wszystkim według niektórych autorów⁹⁴ rozwiązuje problem Fregego–Geacha⁹⁵. Zdaniem Paula Horwicha⁹⁶, argument ten nie jest przeszkodą dla kogoś, kto akceptuje minimalną teorię prawdy. Nie jest on niespójny nawet z emotywizmem, pod warunkiem, że emotywizmu nie będziemy definiowali jako stanowiska, które odrzuca podatność na prawdę⁹⁷. Jeśli prawa pojęciowe dotyczące funkcjonowania predykatu prawdy (takie jak zasada negacji, *modus ponens*, itd.) można byłoby spójnie zinterpretować przy użyciu pojęć prawdy minimalnej i minimalnych stwierdzeń, Hare mógłby rozwiązać problem Fregego–Geacha, nie tylko odwołując się do subatomowych cząstek logiki, lecz również do minimalnej teorii prawdy. Co więcej, mógłby on postulować istnienie, odmiennych od empirycznych, **normatywnych stwierdzeń i normatywnych faktów**⁹⁸.

Choć minimalizm w podejściu do prawdy i przekonań nie przesądza z góry o zwycięstwie nondeskrytywistów w batalii o spełnienie najważniejszych kryteriów dobrej teorii metaetycznej, jednak stanowi dla nich dużą szansę. Dyskusje na temat podatności na prawdę sądów etycznych toczą się obecnie bardzo intensywnie i trudno rozsądzić, która z opcji okaże się bardziej uzasadniona⁹⁹. Hare, powołując się w ostatnim okresie swojej działalności na koncepcję C. Wrighta, stał się w jakimś stopniu zależny od wyniku tej debaty.

Część III Etyka normatywna

Basil Blackwell, Oxford 1990) zalicza się do deflacionizmu, a minimalną teorię prawdy Wrighta traktuje się jako koncepcję wykraczającą poza deflacionizm. Dla Horwicha minimalna teoria prawdy podobna jest bardziej do teorii dyskwo-tacyjnej, Wright natomiast próbuje zbudować najsłabsze pojęcie prawdy, które byłoby spójne z realizmem: nadaje mu oprócz cech odcudzosławiania również cechy normatywne.

⁹⁴ Zob. H. Price, *Semantic...*, dz. cyt., s. 12.

⁹⁵ Zob. § 4.7.

⁹⁶ P. Horwich, *Gibbard's...*, dz. cyt., s. 67-78.

⁹⁷ Zob. P. Horwich, *Truth...*, dz. cyt., s. 87-88.

⁹⁸ Zob. H. Price, *Semantic...*, dz. cyt., s. 12.

⁹⁹ Od lat dziewięćdziesiątych istnieje szeroka dyskusja dotycząca relacji nonkognitywizmu i deflacionizmu. Zastosowanie deflacyjnej czy minimalnej koncepcji prawdy w teoriach nonkognitywistycznych krytykował m.in. P. Boghossian, *The Status of Content*, „Philosophical Review” (94) 1990, s. 157-184; R. Dworkin, *Objectivity and Truth: You'd better Believe It*, „Philosophy and Public Affairs” (2) 1996, s. 87-139.

Część III
Etyka normatywna

Jeśli w Rosji, Chinach czy innych miejscach istniałaby lepsza filozofia, nie musiałoby być dysydentów, a jedynie legalna, rozpoznawalna opozycja czy wielość punktów widzenia wyrażana w wolny sposób; nie musiałoby być również czołgów na placu Tienanmen, ani może również na Bliskim Wschodzie¹.

Choć początki kariery Richarda Hare'a wiążą się, na co wskazuje motto części II, z fascynacją kwestiami metaetycznymi, metaetykę traktował on jedynie jako bardzo istotne pole przygotowujące go do podjęcia głównego celu zainteresowań: racjonalnego rozwiązywania najważniejszych, praktycznych problemów moralnych. Był przekonany, że aby osiągnąć ten cel, potrzebne jest efektywne narzędzie: adekwatny rodzaj normatywnej teorii etycznej. Odpowiadając na to zapotrzebowanie, zaproponował własną koncepcję namysłu moralnego, z której korzystał do końca swoich dni, starając się rozwikłać konkretne, trudne problemy moralne.

Hare w *Myśleniu moralnym* podzielił normatywny namysł moralny na dwie płaszczyzny: intuicyjną i krytyczną. Jego zdaniem, normatywna teoria moralna powinna się rozpaść na odrębne, choć dopełniające się dziedziny rozważań: celem pierwszej jest poszukiwanie idealnej, wzorcowej metody myślenia moralnego, drugiej zaś odpowiedź na pytania, na ile i w jaki sposób człowiek, jako istota poznawczo i psychicznie ograniczona, powinien korzystać z tego wzorca. W związku z tym jego etykę opisujemy, dzieląc ją również na dwie części. Odpowiadać one będą kolejnym rozdziałom części III. W pierwszym z nich (§ 6) przeanalizujemy procedurę idealnego myślenia moralnego, w drugim zaś (§ 7) jej odniesienie do człowieka. Ponieważ zdaniem Hare'a wła-

¹ R. M. Hare, *Philosophy and Conflict*, w: tenże, *Objective Prescriptions...*, dz. cyt., s. 56.

ściwa teoria metaetyczna może mieć bardzo istotny wpływ na rozwój etyki normatywnej, niektóre rozważania przedstawione w części II będą kontynuowane na dalszych stronach niniejszej rozprawy. W części III w znacznej mierze skupimy się jednak na opisie i ocenie Hare'a utylitarystycznej teorii myślenia moralnego.

§ 6. Archanioł – wzór myślenia moralnego

§ 6.1 Dwie płaszczyzny myślenia moralnego

Zdaniem Richarda Hare'a, sposób myślenia moralnego zależy od płaszczyzny, na której się odbywa. Istnieją dwie odrębne, lecz dopełniające się metody: myślenie krytyczne i intuicyjne. Korespondują z nimi dwa rodzaje reguł moralnych: zasady krytyczne są wynikiem myślenia krytycznego, a zasady domniemane rezultatem **myślenia intuicyjnego**². Typowym przykładem bardzo efektywnego zbioru nieskomplikowanych **zasad domniemanych**, których uczymy się w procesie socjalizacji, jest Dekalog. W przeciwieństwie do zasad myślenia krytycznego, prawa tam zawarte są ogólne i proste. Jednak owa prostota powoduje, że w pewnych sytuacjach mogą popadać ze sobą w konflikt. Z tego powodu, choć są uniwersalnymi zaleceniami, nie mają charakteru nadrzędności. Są zasadami *prima facie* – w szczególnych wypadkach ich naruszenie jest uzasadnione³. Zinternalizowane podczas procesu socjalizacji, przybierają charakter prawd dostępnych intuicyjnie. Ponieważ znaczenie deskryptywne zasad domniemanych jest bardzo ustabilizowane, te z nich, które są najszerzej akceptowane, traktujemy jako prawdziwe. To właśnie zasady domniemane opisują deskryptywiści, tworząc naturalistyczne lub nonnaturalistyczne systemy. To w odniesieniu do nich oczywiste wydają się

² Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 80.

³ Zob. tamże, s. 78-79.

takie własności, jak faktualność, prawdziwość i obiektywność sądów moralnych. Jednak najlepszy zbiór reguł domniemanych

to ten, którego przyjęcie owocuje działaniami, dyspozycjami itp. najbardziej zbliżonymi do tych, które zostałyby wybrane, gdybyśmy byli w stanie przez cały czas myśleć krytycznie⁴.

Istnienie zasad domniemanych wynika bowiem z rozpowszechniania i upraszczania zasad krytycznych. W świetle myślenia krytycznego podlegają one również korekcje.

Myślenie krytyczne, wyznaczające pierwszą, nadrzędną płaszczyznę myślenia moralnego, stanowi metodę idealną i wzór, do jakiego powinniśmy dążyć. W tym miejscu powiedzieć można, że jest ono rozumowaniem, którego wynikiem jest wyprowadzenie zasady krytycznej lub wybranie zasady domniemanej spośród tych, które są ze sobą w konflikcie. Polega ono na przeprowadzaniu wstępnych zaleceń, proponowanych przez nas reguł zachowania, przez proces uniwersalizacji w taki sposób, aby w konsekwencji wyselekcjonować uniwersalną i nadrzędną preskrypcję, biorącą pod uwagę istotne preferencje hipotetycznych osób, które są zaangażowane w rozważaną sytuację. Istnieje pewna analogia pomiędzy myśleniem krytycznym a tzw. złotą regułą (wszystko, co byście chcieli, żeby wam ludzie czynili, i wy im czyńcie) czy imperatywem kategorycznym Kanta, choć są między nimi również ważne różnice.

Myślenie krytyczne stanowi metapoziom myślenia intuicyjnego: jest wobec niego epistemicznie i logicznie pierwotne. Uzasadnia sądy domniemane oraz pozwala rozstrzygać konflikty między nimi⁵. Służy człowiekowi do podejmowania decyzji czy oceniania działań w sytuacjach, w których: ma on uzasadnione przekonanie, że konkretne zastane zasady domniemane wchodzą z sobą w konflikt (np. powinność ochrony życia ludzkiego a prawo kobiet do samostanowienia w wypadku aborcji czy powinność niesienia pomocy i zakaz zabójstwa w przypadku eutanazji); nie istnieje zasada domniemana regulująca dany przypadek (np. problemy moralne związane z klonowaniem czy inne współczesne problemy bioetyczne); mamy uzasadnione wątpliwości co do słuszności zastosowania danej zasady domniemanej do konkretnego przypadku.

Dla Richarda Hare'a to, że myślenie krytyczne jest epistemicznie bardziej pierwotne od zasad domniemanych, tzn. może stanowić

⁴ Tamże, s. 68.

⁵ Zob. tamże, s. 40.

narzędzie ich krytyki i zmiany, spowodowane jest między innymi faktem, iż ma ono istotne uzasadnienie teoretyczne, tzn. ufundowane jest na metaetycznej wiedzy i założeniach dotyczących racjonalnego wyboru oraz, poprawnie przeprowadzone, na pytania moralne dostarcza zawsze jednoznacznej odpowiedzi, którą można nazwać obiektywnie prawdziwą. Pozbawione jest dzięki temu podstawowej wady zasad domniemanych – wzajemnego wpadania w konflikt, a przez to pozostawiania człowieka bez moralnej rady. Zdaniem Hare'a, nie jest bowiem prawdą, iż istnieją autentyczne, nierozstrzygalne konflikty wartości:

ci, którzy po prostu stwierdzają, że może zaistnieć nierozstrzygalny konflikt obowiązków, pozostają w swoim myśleniu moralnym na płaszczyźnie intuicyjnej. Na tej płaszczyźnie ów konflikt jest rzeczywiście nierozstrzygalny; ale na płaszczyźnie krytycznej pojawia się warunek, że konflikt rozstrzygniemy, chyba że uznamy, iż nasze myślenie było niepełne. [...] Gdy przeżywasz konflikt obowiązków, jeden z tych obowiązków nie jest twój⁶.

Powoływanie się na występowanie w niektórych bardzo trudnych sytuacjach, w których każda z możliwych do wybrania opcji powoduje odczuwalne zło, „wyrzutów sumienia”, które pojawiają się bez względu na podjętą decyzję, nie jest, według Hare'a, argumentem za istnieniem nierozstrzygalnych konfliktów moralnych. Owszem, są sytuacje, w których bez względu na dokonany przez nas wybór normalny człowiek będzie odczuwał psychiczny ból. Mówiąc o wyrzutach sumienia, często miesza się jednak poczucie żalu, przykrości z poczuciem winy. Możemy być nam bowiem przykro również przy podejmowaniu słusznych decyzji, np. gdy jesteśmy zmuszeni ukarać swoje dziecko. Poczucie winy jest natomiast syntezą przykrości oraz przekonania, że postąpiło się niesłusznie. Przekonanie to jednak musi zostać rozstrzygnięte na innej drodze niż odwołanie się do przykrości. W przeciwnym razie popadniemy w błędne koło. Ową inną drogą jest metoda myślenia krytycznego, dla której kategoria wyrzutów sumienia nie jest istotna.

Hare nie wierzył w autonomiczny zmysł moralny – sumienie. Jest to kolejny ważny powód nadrzędności myślenia krytycznego nad intuicyjnym. Według niego, zasady domniemane, które odkrywamy właśnie poprzez intuicję czy sumienie, są, generalnie rzecz ujmując, jedynie zinternalizowanymi społecznymi regułami zachowania,

⁶ Tamże, s. 40-41.

które nabyliśmy podczas socjalizacji: w domu, na podwórku, w szkole. Są one uproszczonymi zasadami krytycznymi, przekazywanymi z pokolenia na pokolenie. Aby język moralności mógł pełnić swoją funkcję, prawa, które są w nim wyrażane, muszą być bowiem stosunkowo proste, tak by mogły być zapamiętane przede wszystkim przez dzieci. Z powodu tej prostoty popadają one jednak w nierozstrzygalne na poziomie intuicyjnym konflikty.

Takie konflikty nie są udziałem zasad krytycznych, w przeciwieństwie bowiem do domniemanych, mogą być one bardzo szczegółowe i złożone: mogą zawierać w sobie bardzo drobiazgowo i obarczone różnymi warunkami specyfikacje sytuacji, w której ktoś powinien dokonać zalecanego czynu. W pewnych sytuacjach złożoność zasad krytycznych może być pojęta i zapamiętana jedynie przez idealnego nakazodawcę⁷. Chcąc podać przykład znajdowania się w idealnych warunkach epistemicznych, które są istotnym elementem wzorcowej, niezbędnej do pewnego i jednoznacznego ocenienia danej rzeczy, krytycznej procedury namysłu moralnego, Hare przedstawił postać idealnego nakazodawcy – archaniola. **Archaniot** jest bytem idealnym, który nie podlega żadnym ludzkim ułomnościom: ma nadludzkie zdolności myślowe, nadludzką wiedzę, umiejętność przewidywania – przypomina „jasnowidza”. Działa zawsze racjonalnie i bezstronnie, konsekwentnie przestrzegając logiki moralnej gry językowej⁸. Postawiony w obliczu nowej sytuacji będzie w stanie od razu dostrzec wszystkie jej aspekty, w tym następstwa alternatywnych działań, i przedstawić uniwersalną zasadę (może bardzo szczegółową), zgodnie z którą będzie działał, bez względu na to, jaką rolę przyjdzie mu w niej pełnić⁹. Jest on bytem idealnym, stworzonym, aby zobrazować efektywne i poprawne myślenie krytyczne. Postać ta jest kluczowa dla myślenia krytycznego, tak więc stanie się ona głównym przedmiotem naszych dociekań w niniejszym rozdziale.

§ 6.2 Od metaetyki do myślenia krytycznego

Jedną z najważniejszych cech filozofii moralności Richarda Hare'a jest bardzo mocne powiązanie reguł myślenia moralnego z bronio-

⁷ Tamże, s. 142; tenże, *Freedom...*, dz. cyt., s. 39.

⁸ Zob. tamże, s. 61-62.

⁹ Tamże.

nymi przez niego tezami metaetycznymi. Dotyczy to przede wszystkim idealnej, archanielskiej metody myślenia krytycznego. Zdaniem Hare'a, w wielu punktach może być ona wyjaśniona i uzasadniona logiką moralnej gry językowej. Z tego względu warto rozpocząć jej dokładny opis w porządku zależności od tez metaetycznych.

Hare, wyjaśniając strukturę swojej teorii oraz relacje istniejące między metaetyką a etyką normatywną, stwierdził:

na szczycie znajdują się uniwersalne zasady moralne i [...] nie są one stałe, lecz zmienne. [...] Na dole są pojedyncze nakazy działań. One także są zmienne. Ale zakładając, że fakty są twarde i logika stabilna [...], pomiędzy tymi płaszczyznami nie ma dowolności. Nie może być dowolności, gdyż oba końce łączy logika¹⁰.

Analizując treści zawarte w jego pracach z perspektywy historycznej, należy stwierdzić, że swój system budował „z dołu do góry”¹¹. Najpierw zaprezentował w *The Language of Morals* metaetykę, następnie we *Freedom and Reason* i *Myśleniu moralnym* etykę normatywną. Często również dawał wyraz przekonaniu, że etyka normatywna powinna być poprzedzona analizą logiczną i pojęciową oraz na nich się opierać¹². Jednakże owo „opieranie się” można rozumieć wielorako, co zrodziło pewne nieporozumienia.

Zdaniem Bernarda Williama system Hare'a można zrekonstruować w następujący sposób¹³:

(1) Język moralny ma pewne konieczne własności:

(a) z konieczności jest preskryptywny,

(b) uniwersalny,

(c) cechy te są wystarczające, aby zdeterminować myślenie moralne na podstawowym poziomie.

(2) Myślenie krytyczne jest metodą wymagająca utożsamienia się z preferencjami wszystkich osób zaangażowanych w sytuację etyczną.

Zatem:

(3) Myślenie krytyczne jest quasi-pierwszoosobowym namysłem, biorącym pod uwagę wszystkie preferencje zaangażowanych osób, którym rządzi kryteria teorii decyzji.

¹⁰ R. M. Hare, *Foundationalism...*, dz. cyt., s. 117-118.

¹¹ Zob. tenże, *Comments on Williams*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 290.

¹² Zob. tenże, *Myślenie...*, dz. cyt., s. 14.

¹³ Por. B. Williams, *The Structure...*, dz. cyt., s. 185-186.

Powyższa rekonstrukcja skłoniła B. Williamsa do postawienia ważnego pytania: jaki jest status relacji istniejącej między metaetyką (1) a etyką normatywną i myśleniem krytycznym (3), jeśli rozumowanie Hare’a opiera się na dwóch założeniach: metaetyka jest bardziej fundamentalna od etyki normatywnej oraz analiza logiki języka moralnego może być narzędziem krytyki alternatywnych teorii normatywnych lub fundamentem pod budowę własnej koncepcji¹⁴? Czy postulowanie konkretnej metody myślenia moralnego na podstawie przesłanek czysto opisowych, tj. cech języka, nie jest pewnego rodzaju *błędem naturalistycznym*, nieuprawnionym przejściem od formalnych tez dotyczących języka, do tez materialnych mówiących o kryteriach powinności i dobra¹⁵?

B. Williams pojął strukturę systemu Hare’a jako próbę dedukcyjnego wyprowadzenia etyki normatywnej z metaetyki¹⁶. Jednak nie jest to interpretacja trafna, choć winę za to ponosi sam Hare. W różnych swoich tekstach w odmienny sposób rozkładał on bowiem akcenty dotyczące relacji tez metaetycznych do etycznych. W niektórych artykułach daje wyraz daleko idącemu optymizmowi co do możliwości wyprowadzania etyki normatywnej z tez metaetycznych¹⁷. Z drugiej jednak strony, w większości jego głównych prac znajdują się fragmenty, które jednoznacznie stwierdzają niewystarczalność samej analizy pojęciowej do uzasadnienia sposobu rozwiązywania problemów moralnych¹⁸. Wynikiem analizy znaczeń wypowiedzi moralnych nie są bowiem fakty, treści materialne, konkretne jednostkowe czy ogólne oceny i zasady, nie jest nią nawet metoda myślenia krytycznego, lecz jedynie formalne reguły gry językowej – logika moralności, która jest tylko fundamentem uzasadniającym formalną metodę myślenia moralnego. Logika moralności jest jedynie czystą, wyabstrahowaną strukturą języka i jako taka nie może,

¹⁴ Por. tamże.

¹⁵ Zob. tamże, s. 195-196; H. M. Robinson, *Is Hare...*, dz. cyt., s. 73.

¹⁶ Podobną interpretację zależności metaetyki i utilitaryzmu można znaleźć w: W. G. Lycan, *Hare, Singer nad Gewirth on Universalizability*, „The Philosophical Quarterly” (75) 1969, s. 135-144; R. K. Fullinwider, *Fanaticism and Hare’s Moral Theory*, „Ethics” (2) 1977, s. 165-173; w: J. W. Roxbee Cox, *From Universal Prescriptivism to Utilitarianism*, „The Philosophical Quarterly” (142) 1986.

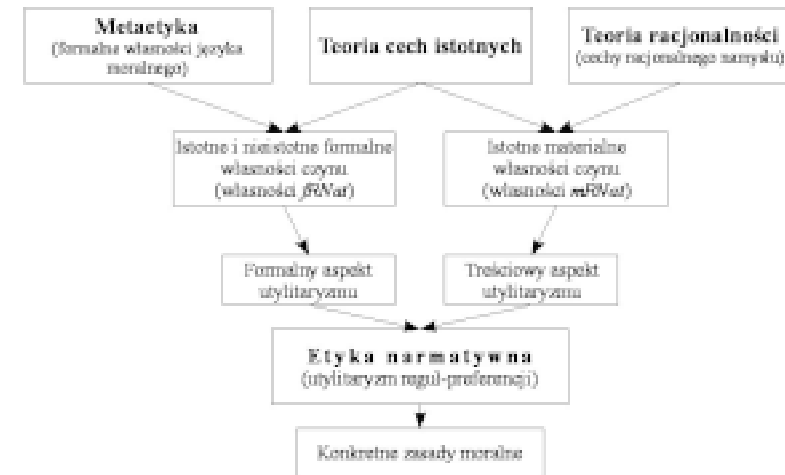
¹⁷ Zob. tenże, *Relevance...*, dz. cyt., s. 209; tenże, *Ethical Theory...*, dz. cyt., s. 215.

¹⁸ Tenże, *Freedom...*, dz. cyt., s. 92-94; tenże, *Myślenie...*, dz. cyt., s. 197; tenże, *Internalism...*, dz. cyt., s. 100.

bez przesłanek dodatkowych, doprowadzić do materialnych wniosków – ani do tego, czy coś jest słuszne, czy nie, ani do tego, jaką metodą to rozstrzygnąć. Hare nie wyprowadził więc etyki normatywnej z samej teorii znaczenia¹⁹; założył również pewną teorię racjonalności oraz teorię cech moralnie istotnych. Podzielał wątpliwość²⁰, że same kryteria formalne mogą doprowadzić do treści materialnych²¹.

Uwzględniając samą logikę, możemy dotrzeć bardzo daleko – dalej, niż wielu sądziło – z tym jednak, że jeżeli chodzi o wybór zasad stosowanych w tym naszym świecie, istotne znaczenie mają fakty dotyczące świata i zamieszkujących go ludzi²².

Relacje występujące pomiędzy najważniejszymi elementami jego systemu, które analizować będziemy na dalszych stronach niniejszej części III, można przedstawić w następujący sposób:



Rysunek 8. Struktura teorii R. M. Hare’a

¹⁹ Zob. tenże, *Utilitarianism and the Vicarious Affects*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 235.

²⁰ Wątpliwość taką przedstawiła np. G.E.M. Anscombe. Zob. G.E.M. Anscombe, *Modern Moral Philosophy*, w: W. D. Hudson (ed.), *The Is-Ought...*, dz. cyt., s. 177.

²¹ Zob. R. M. Hare, *Comments on Nagel*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 249.

²² Tenże, *Myślenie...*, dz. cyt., s. 16.

§ 6.3 Myślenie krytyczne:
teoria cech istotnych i proces ich uniwersalizacji

W celu przeprowadzenie namysłu moralnego musimy dysponować wiedzą o faktach. Nie wszystkie jednak fakty są moralnie istotne. Kryteria wyboru spośród istotnych dla namysłu moralnego cech naturalnych wraz z ich uzasadnieniem można nazwać **teorią cech istotnych**. W świetle systemu Hare'a teoria naturalnych cech moralnie relewantnych (*RNat*) ma aspekt formalny i materialny. Własności istotne ze względu na proces poznawczy, który do nich prowadzi, można podzielić na **istotne formalnie** (*fRNat*) oraz **materialnie** (*mRNat*). Hare nie przedstawił powyższej dychotomii, ułatwi nam ona jednak dalszą analizę. To, jakie własności naturalne powinniśmy wziąć pod uwagę, które z cech naturalnych są moralnie relewantne, możemy wywieść jedynie częściowo z analizy logiki języka moralnego. Ważnym elementem są również przesłanki materialne. Moralnie istotne własności formalne wyznaczamy, badając nasze intuicje lingwistyczne oraz konsekwencje wynikające z formalnych praw moralnej gry językowej; poznanie istotnych własności materialnych wymaga pewnych dodatkowych założeń oraz wiedzy na temat ocenianej, konkretnej sytuacji etycznej. Abyśmy mogli określić, które własności *mRNat* są moralnie relewantne, nie wystarczy znać natury ludzkiej czy własności języka moralnego. W procesie poznawania własności *mRNat* rozum praktyczny i doświadczenie są ważniejsze od rozumu teoretycznego²³. Jeśli wybór formalnych cech *fRNat* jest od procesu oceny czynu niezależny, to taka zależność zachodzi w przypadku wyznaczania cech materialnych. Poszukiwanie własności *mRNat* nie sprowadza się bowiem do analizy znaczenia czy ustalenia warunków prawdziwości. To, że powinniśmy poddawać uniwersalizacji te a nie inne cechy materialne, możemy poznać jedynie hermeneutycznie: stosowanie zasad moralnych wyznacza własności istotne materialnie, a my, zakładając, co jest istotne, tworzymy zasady²⁴.

Ujęcie jakiejś cechy działania czy sytuacji jako moralnie relewantnej oznacza zastosowanie do tego działania lub sytuacji zasady moralnej, która o tej cesze wspomina. [...] Początkowo musimy więc zdać się na domysły. Jeśli uważamy, że jakaś cecha sytuacji *mogłaby* być relewantna, eksperymentujemy zasa-

²³ Zob. tamże, s. 82-85.

²⁴ Zob. tamże, s. 83-84.

dami wspominającymi o tej cesze; zaakceptowanie zasad będzie równoznaczne z zaakceptowaniem relewancji tej cechy [...] Błędem byłoby sądzić, że można *najpierw* wybrać moralnie istotną cechę sytuacji, a dopiero *później* pytać, jakie zasady moralne należy do niej zastosować. To właśnie zasady określają to, co jest relewantne²⁵.

Wyznaczenie materialnych cech moralnie relewantnych (*mRNat*) w dużej mierze zależy od procesu myślenia krytycznego, polega on bowiem na uniwersalizowaniu wstępnych zaleceń na podstawie wybranych cech moralnie istotnych. Uniwersalizacja natomiast, jeśli weźmiemy pod uwagę odpowiednie własności relewantne, stanowi istotę myślenia krytycznego.

Jeśli uniwersalność (*universality*) jest cechą sądów moralnych, a uniwersalizowalność (*universalisability*) cechą zaleceń polegającą na możliwości uczynienia z nich w pełni uniwersalnych sądów moralnych, to **uniwersalizacja** (*to universalize*) w świetle teorii Hare'a jest praktyczną metodą osiągnięcia owej uniwersalności. Uniwersalizacja to proces, którego wynikiem jest w pełni uniwersalna zasada moralna, argumentami zaś są zalecenia oraz istotne dla moralności cechy naturalne (*RNat*), które przy ocenie muszą być wzięte pod uwagę; jest to proces przemiany imperatywów jednostkowych, uniwersalnych niepełnie, sądów quasi-moralnych, nie w pełni uniwersalnych („Tylko ja i tylko w tej sytuacji powinienem tylko tobie zrobić x”) w pełną zasadę moralną. Podczas uniwersalizacji istotne cechy *RNat* wprowadzamy do prawa moralnego i wiążemy kwantyfikatorem ogólnym²⁶.

Według J. L. Mackiego w etyce Hare'a wyróżnić można trzy stopnie uniwersalizacji²⁷. Polegają one na:

- I. Eliminacji różnic numerycznych.
- II. Postawieniu siebie na miejscu innych osób;
- III. Wzięciu pod uwagę różnych upodobań i konkurencyjnych idei.

R. M. Hare nie zgodził się jednak na powyższy podział²⁸. Jego zdaniem, J. L. Mackie pomieszał ze sobą uniwersalizowalność jako własność logiczną z moralną bezstronnością, która może być jej

²⁵ Tamże, s. 117-118, 83.

²⁶ Zob. tamże, s. 78-79.

²⁷ Zob. J. L. Mackie, *Ethics...*, dz. cyt., s. 83-102. Komentarz do omawianego podziału znaleźć można w: P. Singer, *Reasoning...*, dz. cyt., s. 147-159.

²⁸ Zob. R. M. Hare, *Comments on Nagel...*, dz. cyt., s. 249.

wynikiem przy uwzględnieniu przesłanek dodatkowych. Za pomocą logiki języka moralnego możemy bowiem uzasadnić co najwyżej te aspekty myślenia krytycznego, które Mackie nazwał pierwszym i drugim etapem uniwersalizacji²⁹.

Sądy moralne uniwersalizują się [...] tylko w jednym sensie, w tym mianowicie, że pociągają za sobą identyczne sądy na temat wszystkich przypadków tożsamych pod względem swoich własności uniwersalnych³⁰.

Pomimo iż Hare nie uznał rozróżnienia przeprowadzonego przez Mackiego, trafnie oddaje ono jednak kolejne etapy krytycznego namysłu moralnego wchodzącego w skład jego etyki normatywnej. Polegają one na uniwersalizacji zaleceń przy uwzględnieniu kolejnych zbiorów istotnych cech *RNat*. Z tego też względu, niezależnie od tego, że błędem jest mówienie o trzech rodzajach uniwersalizacji, analogiczna hierarchia rodzajów własności istotnych *RNat* oraz przyporządkowanych im etapów myślenia krytycznego wydaje się przydatna.

§ 6.3.1 Pierwszy etap myślenia krytycznego

Zdaniem Hare'a, pierwszy oraz drugi etap myślenia krytycznego (procesu uniwersalizacji) ma charakter formalny. Polega on na związaniu kwantyfikatorem ogólnym istotnych formalnie cech sytuacji etycznej (*fRNat*). Aby dowiedzieć się, które własności należy wziąć pod uwagę, niezbędne jest wyciągnięcie wniosków z formalnych własności języka moralnego. Taka analiza dostarczyć nam może wiedzy negatywnej oraz pozytywnej: informacji, jakich własności nie powinniśmy, a jakie musimy brać pod uwagę przy dokonywaniu ocen moralnych. Cechy *nie-fRNat*, które nie mogą być istotne dla dyskursu moralnego, podobnie jak wszystkie te cechy *fRNat*, które musimy brać pod uwagę, można wyprowadzić z uniwersalności i preskryptywności sądów moralnych³¹.

Wymóg uniwersalności języka moralnego jest odpowiedzialny za **pierwszy etap myślenia krytycznego**. Konsekwencje wynikające z prawa superweniencji wskazują przede wszystkim, jakich cech *nie-*

²⁹ Zob. tamże.

³⁰ Tenże, *Myślenie...*, dz. cyt., s. 138.

³¹ Zob. tenże, *Relevance*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 209; tenże, *Ethical Theory and Utilitarianism*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 215.

fRNat nie możemy traktować jako moralnie istotnych. Na drodze analizy wymogu uniwersalności możemy dowiedzieć się bowiem, że nie powinniśmy zważać na własności, które nie są uniwersalne, jak również, że dla oceny istotne jest branie pod uwagę faktów³². Jeśli sądy moralne mają być uniwersalne, nie mogą być one wydawane z użyciem stałych indywidualnych: ludzie, sytuacje i czyny powinny być opisywane jedynie za pomocą predykatów odnoszących się do własności uniwersalnych. Wyklucza to z dziedziny moralności te preskrypcje, które są wyrazem pewnego rodzaju egoizmu, np. stanowisko, które zezwala na wydanie sądu dotyczącego wszystkich, z wyjątkiem Krzysztofa Saji.

Jeśli nie zmienia się sprawca danego czynu ani relacje pomiędzy ludźmi uwikłanymi w daną sytuację, ani ich przekonania i postawy, zmianie ulegnie natomiast czas i przestrzeń, w których czyn zachodzi lub w których dokonujemy oceny, to, zdaniem Hare'a, ocena moralna wydana w czasoprzestrzeni t_1 lub na temat sytuacji dziejącej się w t_1 musi być taka sama jak ocena wydana w czasoprzestrzeni t_2 lub na temat działania dokonanego w czasoprzestrzeni t_2 . Ponieważ punkty czasowy i przestrzenny nie są własnościami uniwersalnymi, lecz indywidualnymi, nie są one moralnie relewantne³³. Co więcej, jeśli zmienia się czas i przestrzeń oraz sprawca danego czynu z osoby O_1 na osobę O_2 , nie ulegną natomiast zmianie istotne cechy psychiczne, fizyczne i społeczne sprawcy, odbiorców działania ani świata (nie ulegną zmianie inne istotne własności sytuacji etycznej oraz zaangażowanych osób³⁴), z konieczności ocena wydana na temat działania osoby O_1 oraz O_2 musi być taka sama. Ponieważ uniwersalność wymusza na nas nietraktowanie osób jako indywidualów, a jedynie jako podmioty opisywane za pomocą predykatów uniwersalnych, zatem osoby jako indywiduala nie są relewantne dla oceny dokonanych przez nie czynów³⁵.

³² Zob. tenże, *Relevance...*, dz. cyt., s. 207-208.

³³ Zob. tenże, *Myślenie...*, dz. cyt., s. 131, 146-147.

³⁴ Przez **zaangażowane osoby** będziemy rozumieć wszystkie te osoby, na które w istotny sposób wpływa obowiązywanie danego prawa moralnego, w których interesie leży taki a nie inny kształt rozważanej zasady moralnej.

³⁵ Zob. tamże, s. 83. Ze stanowiskiem tym polemizuje D. Locke, twierdząc, iż teza, że względ na osoby nie może być istotny w namyśle moralnym, sama ma charakter etyczny. Zob. D. Locke, *The Trivializability of Universalizability*, „The Philosophical Review” (1) 1968, s. 37.

Wymóg abstrahowania od takich własności ocenianej sytuacji, których nie można opisać za pomocą predykatów uniwersalnych, powoduje, że utilitaryzm Hare'a jest teorią uniwersalistyczną. W konsekwencji sądy moralne posiadają ważną zaletę: są zabezpieczone przed indywidualizmem, w który, zdaniem Hare'a, popada, np. partykularyzm J. P. Sartre'a³⁶. Nie tracą one pewnej dozy ogólności, rozciągają się bowiem na wiele podobnych czynów. Dzięki temu zasady oceniające stają się „zasadami wielokrotnego użytku”, co powoduje, że mają one wartość poznawczą i mogą służyć wszystkim, którzy znaleźli się w podobnej sytuacji³⁷.

§ 6.3.2 Drugi etap myślenia krytycznego

Nieistotność czasu, przestrzeni oraz indywidualnych cech osób faktycznie można wywieść z uniwersalności. Jednak aby dany sąd moralny nie był względny wobec indywidualnych preferencji i przekonań, nie możemy ograniczać się jedynie wymogami prawa superwencji HSup I³⁸. Aby uzyskać większą dozę bezstronności, należy przeprowadzić również **drugi etap myślenia krytycznego**. Wyprowadzając jakiś nakaz, musimy wydać go jako obowiązujący we wszystkich sytuacjach, które są identyczne pod względem moralnie istotnych własności oraz bez względu na rolę, jaką mielibyśmy w niej przyjąć (lub która byłaby nam w niej narzucona)³⁹. Polega to na postawieniu się na miejscu innych osób oraz odpowiedzeniu sobie na pytanie, czy również patrząc z tej perspektywy, chcielibyśmy, aby

³⁶ Krytykę indywidualizmu J. P. Sartre'a, przedstawionego np. w: J. P. Sartre, *Egzystencjalizm jest humanizmem*, tłum. J. Krajewski, KiW, Warszawa 1957, znaleźć można w: R. M. Hare, *Universalisability...*, dz. cyt., s. 21-28; tenże, *Freedom...*, dz. cyt., s. 41, lub tenże, *Uniwersalny Preskrytywizm*, w: P. Singer (red.), *Przewodnik...*, s. 505. Z drugiej strony, niektórzy autorzy zarzucili Hare'owi, iż zasada uniwersalizowalności jest błędna, gdyż nie docenia ona indywidualnego charakteru wielu ocen, które nie podlegają prawu superwencji. Zob. A. MacIntyre, *What Morality Is Not*, „Philosophy” (32) 1957, s. 332-333; P. Winch, *The Universalizability of Moral Judgments*, „Monist” (49) 1965, s. 196-214; E. W. James, *Working in and Working to Principles: Penn's Lie nad Hare's Myth of Universalizability*, „Ethics” (83) 1972, s. 51-57. Obronę partykularyzmu w etyce odnaleźć można w: J. Dancy, *Ethics without principles*, Oxford University Press, Oxford 2004.

³⁷ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 51-52.

³⁸ Zob. s. 159.

³⁹ Tenże, *Sorting...*, dz. cyt., s. 25.

obowiązywało wydane przez nas pierwotne zalecenie. Należy wziąć pod uwagę także taką hipotetyczną sytuację, w której role między osobami interakcji zostaną odwrócone⁴⁰. Wygłaszając preskrypcję, jako osoby znajdujące się w pewnej sytuacji, musimy pozytywnie odpowiedzieć na pytanie, czy chcielibyśmy, aby ten nakaz obowiązywał również wtedy, gdy bylibyśmy na miejscu innej osoby zaangażowanej w daną sytuację etyczną, a ona, będąc na naszym miejscu, działałaby wedle postulowanej zasady. Na przykład jeśli chcielibyśmy wydać jako pacjentowi określone zalecenie, musielibyśmy przeprowadzić eksperyment myślowy, czy znajdując się w podobnej sytuacji jako lekarz czy pielęgniarka, również pragnęlibyśmy, aby wydane przez nas zalecenie obowiązywało.

Zatem na drugim etapie myślenia krytycznego do zbioru *RNat* należy włączyć istotne psychofizyczne oraz społeczne cechy posiadane przez podmioty zaangażowane w sytuację etyczną, które są możliwe do opisanego z perspektywy osoby trzeciej, np. dobra pamięć, bogactwo lub bieda, siła fizyczna, kalectwo umysłowe lub fizyczne, wysoki status społeczny, posiadanie rodziny, bycie samotnym itp. Na tym etapie pod uwagę musimy wziąć nie tylko faktyczne możliwości zmiany ról, ale również prawdopodobne sytuacje hipotetyczne. Według Hare'a, taka zuniwersalizowana zasada moralna nie różnicuje już ludzi ze względu na ich określone umiejętności, zdolności, status społeczny itd. Nie unieważnia ona jednak różnic związanych ze smakiem, wyznawanymi ideami i wartościami. Aspekty te grają rolę na kolejnym etapie uniwersalizacji.

Niektórzy filozofowie twierdzą, że uniwersalność jako wynik powyższego etapu myślenia krytycznego nie jest formalną własnością języka. Według Thomasa Nagela, to, co Hare nazywa własnością formalną, jest materialną, moralną zasadą bezstronności⁴¹. Podobną tezę wysunął John L. Mackie⁴². Rozróżniwszy trzy stopnie uniwersalizacji, próbował wykazać, że tylko o pierwszym z nich można orzec, że jest charakterystyczny dla myślenia moralnego w ogóle, pozostałe natomiast są przemyconymi liberalno-protestanckimi poglądami Hare'a⁴³. Hare uważałby jednak, że również powyżej opisa-

⁴⁰ Zob. tenże, *Myślenie...*, dz. cyt., s. 144-145.

⁴¹ Por. T. Nagel, *The Foundations...*, dz. cyt., s. 108-110.

⁴² Zob. J. L. Mackie, *Ethics...*, dz. cyt., s. 83-102.

⁴³ Por. tamże, s. 83-102.

na bezstronność związana z drugim etapem myślenia krytycznego wynika z koniunkcji preskryptywności i uniwersalności. Jego zdaniem,

bezstronność, która jest skutkiem Rawlsa «zasłony niewiedzy» [w uniwersalnym preskryptywizmie – K. S.] uzyskana jest w oparciu o środki czysto formalne. [...] Z tego też względu, rozważając samą logikę pojęć moralnych, możemy przyjąć tak silną pozycję, do jakiej doszedł Rawls, który używał bardziej skomplikowanych narzędzi o słabszych podstawach⁴⁴.

Hare argumentował, że skoro zalecenia moralne muszą być uniwersalne, aby nie być wewnętrznie sprzecznymi, musimy zwracać uwagę na własności posiadane przez oceniane przez nas rzeczy. Nie możemy przypisać innego predykatu moralnego tym samym istotnym cechom naturalnym. Z tego powodu wiedza dotycząca własności naturalnych jest relewantna⁴⁵. To natomiast, że uniwersalność wymusza na nas wzięcie pod uwagę sytuacji hipotetycznych, spowodowane jest faktem, iż rozciąga się ona na każdą istotnie podobną sytuację, która zachodzi lub mogłaby zajść. Ponieważ każdy, kto szczerze wypowiada preskrypcję, musi pragnąć rzeczy, której zaleca⁴⁶, a uniwersalność wymusza na nas nietraktowanie osób jako indywiduów, to aby wydać sąd moralny, musimy odpowiedzieć na pytanie, czy chcielibyśmy, aby wyrażone w nim zalecenie obowiązywało uniwersalnie. Według Hare'a, aby uzyskać odpowiedź na powyższe pytanie, najlepszą metodą jest postawienie się na miejscu innych osób i zbadanie, co odpowiedzielibyśmy, patrząc z ich perspektywy. Ponieważ należy wziąć pod uwagę każdą rolę, którą moglibyśmy pełnić w hipotetycznych sytuacjach, musimy im wszystkim przyznać jednakową wagę. Nie możemy konsekwentnie nadawać większego znaczenia perspektywie, która np. aktualnie nas dotyczy⁴⁷, i umniejszać pozycję, co do której mamy przeświadczenie, że nigdy nie będzie faktycznie naszą (lub gdy takie prawdopodobieństwo jest bardzo małe), gdyż przeprowadzając pierwszy etap myślenia krytycznego, powinniśmy abstrahować od stałych indywiduowych – tego, kim jesteśmy, w jakim czasie i miejscu się znajdujemy. W kon-

⁴⁴ R. M. Hare, *Relevance...*, dz. cyt., s. 209. Porównanie teorii R. M. Hare'a i J. Rawlsa można znaleźć w: M. H. Lessnoff, *Justice, Social Contract, and Universal Prescriptivism*, „Philosophical Quarterly” 110 (1978), s. 65-73.

⁴⁵ Zob. tamże, s. 207-208.

⁴⁶ Tamże, s. 210.

⁴⁷ Zob. tenże, *Ethical Theory and Utilitarianism*, dz. cyt., s. 215.

sekwencji ze względu na to, że wolność naszych nakazów ograniczona jest formalnym wymogiem uniwersalności, nie możemy chcieć obowiązywania pewnych reguł:

a) z powodu niemożności logicznej (gdy, nasze pragnienia uformowane w uniwersalną regułę zniosą możliwość ich zapragnienia);

b) z powodu niemożności praktycznej (gdy, stawiając się na miejscu innych osób, nie chcielibyśmy, jeśli jesteśmy osobami racjonalnymi, aby pewne reguły obowiązywały uniwersalnie⁴⁸).

Gdybyśmy myśląc krytycznie, pozostali jedynie przy pierwszym z powyższych punktów (a), pociągnęłoby to za sobą poważne problemy. Moglibyśmy bowiem zaakceptować uniwersalny imperatyw, który mówi:

„Kłam, łam obietnice, kradnij, oszukuj na podatkach, gdy tylko sytuacja tego wymaga” – pod warunkiem jednak, że sytuacja nie wymaga tego często, a ty masz duże prawdopodobieństwo, że nie zostaniesz przyłapany⁴⁹.

Z punktu widzenia logiki można chcieć takiej zasady, gdyż pomimo jej stosowania instytucje moralne przetrwają, możliwość czynienia wyjątków również. W oczach S. Blackburna powyższy argument dyskwalifikuje imperatyw kategoriowy Kanta⁵⁰. Nie podważa on jednak teorii Hare'a. Hare nie ogranicza się bowiem jedynie do czysto logicznej niemożliwości chcenia danej zasady. Na drugim etapie myślenia krytycznego uniwersalizacja wymaga odrzucenia wstępnego zalecenia również wtedy, gdy jako w pełni racjonalne, roztropne podmioty, patrzące z bezstronnej perspektywy, nie chcielibyśmy, aby ono obowiązywało uniwersalnie. Hare uzasadnia wagę preferencji wszystkich osób, również tych hipotetycznych, uniwersalnością i preskryptywnością. Choć jest prawdą, że nie zawsze wydawane przez nas nakazy muszą mieć źródło w naszych pragnieniach czy preferencjach, trudno jednak pojąć preskrypcję, która nie są związane z żadną preferencją czy interesem. Funkcją preskrypcji jest bowiem wyrażanie preferencji. Drugi etap uniwersalizacji nie opiera się jednak, jak mogłoby się wydawać i jak sądził Hare, jedynie na logicznych własnościach języka moralnego. Zakła-

⁴⁸ Ten punkt związany jest już jednak z trzecim etapem myślenia krytycznego i należy się przychylić do opinii, że nie jest on czysto formalny.

⁴⁹ Zob. S. Blackburn, *Sens dobra. Wprowadzenie do etyki*, przeł. T. Chawziuk, Dom Wydawniczy Rebis, Poznań 2002, s. 185.

⁵⁰ Zob. tamże, s. 185-186.

da on bowiem również, że podmioty moralne powinny być racjonalne w świetle pewnej koncepcji racjonalności.

§ 6.3.2.1 *Fanatyk*

Z drugim etapem myślenia krytycznego związany jest pewien problem. Można sobie wyobrazić fanatycznego i w pełni racjonalnego nazistę, który przeprowadziwszy drugi etap uniwersalizacji, szczerze głosi zalecenie, by zgładzić wszystkich Żydów. Mogłoby się tak zdarzyć, gdyby jego preferencja była na tyle silna, iż gotów byłby on poświęcić dla jej spełnienia swoje życie, nawet gdyby wyobraził sobie, że sam jest ofiarą własnego zalecenia. Innym przykładem jest osoba o upodobaniach sadomasochistycznych, która szczerze zalecałaby zadawanie bólu każdemu, nawet jeśli taka preskrypcja miałaby odnosić się do jej samej. Kolejną egzemplifikacją mógłby być pewny siebie egoista, który uniwersalnie głosiłby: „Każdy powinien dbać tylko o swój własny interes”. Jeśli poprzestalibyśmy na dwóch etapach myślenia moralnego, powyższe zalecenia należałoby nazwać moralnymi.

Argument przywołujący istnienie fanatyków jest jednym z bardziej znanych zarzutów przeciwko uniwersalnemu preskryptywizmowi⁵¹. Hare odpowiedział na niego w dwojaki sposób. Po pierwsze, fanatycy o powyższych cechach istnieją jedynie teoretycznie. Bardzo rzadko, jeśli w ogóle, można spotkać ludzi, którzy przejawiają konsekwentne tendencje do fanatycznego wyznawania pewnych poglądów, tzn. potrafią szczerze zaakceptować uniwersalną preskrypcję o fanatycznych treściach po przeprowadzeniu dwóch etapów myślenia moralnego. Najprawdopodobniej nie istnieje antysemita, który tak mocno chce zgładzić Żydów, że posunąłby się do samobójstwa, gdyby sam okazał się Żydem. Również przedstawiony przypadek sadomasochisty wydaje się być tworem teoretycznym.

Powyzsza linia obrony teorii Hare’a nie do końca jednak przekonuje. Można bowiem twierdzić, że przykłady fanatyków, choć nieczęsto spotykane, istnieją realnie. Spotkać można egoistów, którzy

⁵¹ Zob. A. Gettner, *Hare and Fanaticism*, w: „Ethics” (2) 1977, s. 160-164; R. K. Fullinwider, *Fanaticism...*, dz. cyt. Obronę teorii Hare przed zarzutami postawionymi w powyższych artykułach znaleźć można w: J. Narveson, *Liberalism, Utilitarianism, and Fanaticism: R. M. Hare Defended*, „Ethics” (3) 1978, s. 250-259. R. M. Hare, *Wrong and Harm*, w: *Essays on the Moral...*, dz. cyt., s. 102-108; tenże, *Myślenie...*, dz. cyt., s. 213-233.

nawołują do egoistycznego dbania jedynie o własny interes. Co więcej, dobra teoria moralna nie powinna, nawet dla nietypowych osób, implikować tezy, iż moralność danego zalecenia jest funkcją poświęcenia i ofiary, jakie jesteśmy w stanie ponieść, aby uznać zalecenie za moralne. Skuteczniejszym sposobem odparcia powyższego zarzutu jest wskazanie na niewystarczalność dwóch etapów uniwersalizacji. Odpowiedź druga głosiłaby zatem, że fanatycy czy egoiści nie mogliby zuniwersalizować swoich kontrowersyjnych zaleceń, gdyby przeprowadzili również trzeci etap myślenia krytycznego.

§ 6.3.3 *Trzeci etap myślenia krytycznego*

Przykładowy nazista, nawet jeśli zalecałby eksterminację Żydów, zakładając, że preskrypcja ta odnosić się może również do niego, nie stanowi teoretycznego problemu, gdyż na trzecim etapie myślenia moralnego jego preferencja, której daje wyraz w swoim zbrodniczym zaleceniu, musi zostać skonfrontowana z innymi preferencjami innych osób, niepodzielających jego poglądów. Na tym etapie nie pytamy już: *Czy chciałbyś, aby twoje wstępne zalecenie obowiązywało, gdyby zmieniła się twoja fizyczna, obiektywna sytuacja?* (np. stałbyś się Żydem), lecz: *Czy chciałbyś, aby twoje wstępne zalecenie obowiązywało, gdyby zmieniły się posiadane przez ciebie przekonania, pragnienia, cele lub preferencje – stany, które cię motywują.* W konsekwencji przykładowy sadomasochista, aby skutecznie zuniwersalizować swoją wstępną preferencję: *Niech będzie nieograniczone zadawanie bólu!*, musiałby zapytać siebie, czy chciałby jej obowiązywania również wtedy, gdyby przestał mieć skłonności sadomasochistyczne i tak jak normalny człowiek, traktował ból jako doświadczenie, którego usilnie się unika. Również egoista musiałby postawić się na miejscu osób bardzo mocno pragnących pomocy i zadać sobie pytanie, która z preferencji jest silniejsza: jego, aby nie musieć jej udzielać czy osób bardzo potrzebujących wsparcia. Zdaniem Hare’a, biorąc pod uwagę cechy aktualnego świata i zamieszkujących go istot, nikt racjonalny szczerze nie mógłby zaakceptować fanatycznych czy egoistycznych uniwersalnych zaleceń. Biorąc pod uwagę jedynie bezstronnie ocenioną siłę posiadanych przez różnych ludzi preferencji, nigdy chęć zamordowania Żydów, zadania bólu czy choćby niepomagania innym nie będzie tak silna jak pragnienie

uniknięcia cierpienia i niebycia pozbawionym życia. Można sobie wyobrazić taki możliwy, fanatyczny świat, nie jest on jednak istotny, ponieważ jest mało prawdopodobny.

Zdaniem Hare'a, nie powinniśmy brać pod uwagę preferencji osób będących w niezwykłych sytuacjach hipotetycznych⁵², gdyż zakres osób i ich preferencji, które przy ocenie moralnej musimy wziąć pod uwagę, jest ograniczony wymogiem moralnej relewancji. W literaturze przedmiotu spotkać można bowiem wiele zarzutów skierowanych w stronę kalkulacji utylitarystycznej, które polegają na przedstawieniu hipotetycznej, bardzo mało prawdopodobnej, a wręcz nierzeczywistej sytuacji, zastosowaniu do niej metody utylitarystycznego namysłu, a następnie postawieniu tezy, iż jej konsekwencje są sprzeczne z naszymi intuicjami moralnymi. Hare uważał że sytuacje, których zajście jest bardzo nieprawdopodobne, należy traktować jako nerelewantne. Myślenie krytyczne dobrze spełnia swoją rolę w świecie takim, jakim widzimy go codziennie, może jednak dawać kontrintuicyjne rezultaty, jeśli zastosujemy je do fantastycznych, niespotykanych sytuacji⁵³. Moralnie znaczące są, jakby powiedział R. Nozick, tylko „bliskie światy możliwe”, tzn. takie, które prawdopodobnie mogłyby istnieć⁵⁴. Uzasadnienie powyższego ograniczenia przebiega dwutorowo. Utylitarystyka nie musi się sprawdzać we wszystkich światach możliwych, ponieważ, po pierwsze, metoda myślenia krytycznego nie opiera się jedynie na analitycznych własnościach języka moralnego, lecz również na dodatkowych założeniach materialnych pochodzących ze świata aktualnego. Gdyby metoda utylitarystyczna wynikała z praw pojęciowych, powinna byłaby obowiązywać we wszystkich światach możliwych; ponieważ jednak została wzbogacona prawidłowościami ze świata aktualnego, stawianie jej zarzutu, że nie ma tak szerokiego zasięgu, jest nieporozumieniem. Zrozumiałe jest wymaganie, aby formalne elementy omawianego systemu obowiązywały we wszystkich możliwych światach, nie może ono odnosić się jednak do materialnych założeń związanych z myśleniem krytycznym. Co więcej, powoływanie się na intuicje moralne, które miałyby przeczyć zaleceniom wynikającym z zastosowania myślenia krytycznego, może być zabiegiem nieroztropnym, jeśli takie intuicje miałyby się odnosić do niespotykanych wcześniej przez

⁵² Zob. tamże, s. 143-147, 225.

⁵³ Por. tamże, *Utilitarianism and the Vicarious...*, dz. cyt., s. 235.

⁵⁴ Zob. R. Ziemińska, *Eksternalizm...*, dz. cyt., s. 138.

ludzi sytuacji. Zdaniem bowiem Hare'a, intuicje moralne mają „konserwatywne skłonności” i nie radzą sobie w sytuacjach, w których nie miały szans się rozwijać.

Trzeci etap myślenia krytycznego charakteryzuje się tym, że wiedza potrzebna do jego przeprowadzenia dotyczy subiektywnych stanów mentalnych zaangażowanych osób, a nie, jak na drugim etapie, jedynie ich obiektywnych, zewnętrznych cech. Jeśli typowe własności istotne dla drugiego etapu myślenia krytycznego są ujmowane z „zewnętrznego punktu widzenia”, to relewantne własności naturalne, które musimy wziąć pod uwagę na finalnym etapie namysłu moralnego, są wewnętrznymi, psychicznymi stanami zaangażowanych osób.

Będziemy musieli pamiętać o rozróżnieniu zwykłej wiedzy, że coś się komuś stało, i wiedzy o tym, *co to dla niego oznacza*. I właśnie tę wiedzę drugiego rodzaju powinniśmy traktować jako relewantną oraz niezbędną do uzyskania informacji potrzebnych do racjonalnego formułowania sądów moralnych⁵⁵.

Archaniot, jako idealny podmiot moralny, nie ma trudności ze zdobyciem tej wiedzy. Według Hare'a człowiek, aby ją osiągnąć, powinien zrobić użytek ze znanej wszystkim ludziom metody myślowego i emocjonalnego utożsamiania się z innymi osobami. Pełne utożsamienie się z innymi osobami zakłada natomiast tożsamość z ich stanami motywacyjnymi. Przekonanie to jest istotnym założeniem trzeciego etapu myślenia krytycznego. Nazwać je można **w warunkiem pełnego utożsamienia się z innym**⁵⁶:

Nie mogę poznać skali i jakości cudzych cierpień oraz, ogólnie rzecz biorąc, motywacji i preferencji, jeżeli nie kieruje mną taka sama motywacja w odniesieniu do tego, co mogłoby mnie spotkać, gdybym znalazł się na miejscu kogoś innego wraz z jego motywacjami i preferencjami⁵⁷.

Trzeci etap myślenia krytycznego wymaga od nas wiedzy, którą człowiek zdobyć może poprzez pełne „postawienie się na miejscu innej osoby”⁵⁸. Musimy przyjąć jej istotne dla oceny pragnienia, preferencje, interesy, poglądy i wartości⁵⁹ – istotne, fenomenalne, subiektywne aspekty tej świadomości. Możliwość utożsamienia się z preferencjami in-

⁵⁵ Tenże, *Myślenie...*, dz. cyt., s. 120.

⁵⁶ Krytykę tego warunku wraz z próbą przedstawienia Hare'a uzasadnienia metody myślenia krytycznego można znaleźć w: M. McDermott, *Hare's Argument for Utilitarianism*, „The Philosophical Quarterly” (133) 1983, s. 386-391.

⁵⁷ R. M. Hare, *Myślenie moralne*, s. 139, 129.

⁵⁸ Zob. tamże, s. 120-124, 139.

⁵⁹ Zob. tenże, *Ethical Theory and Utilitarianism*, dz. cyt., s. 219-220.

nych osób stanie się bardziej jasna, jeśli zauważymy, że ludzie często to czynią (choć w niedoskonały sposób) za sprawą empatii, współodczuwania. Osoba, która współodczuwa, nie traktuje innych jako osoby by trzeciej, ale „staje się nimi”, przyjmując subiektywne, fenomenalne aspekty ich świadomości – ich uczucia i pragnienia. Jeśli zatem współodczuwając, doznajemy uczuć innych jako swoich, motywują nas one do działania. Na trzecim etapie myślenia krytycznego najważniejszą więc cnotą wydaje się być empatyczna wrażliwość. Musimy sobie wyobrazić, jak to jest być kimś innym, hipotetyczną osobą zaangażowaną w dane działanie. Ostatecznie wymaga to od nas, abyśmy przy wydawaniu zalecenia moralnego zachowali bezstronność zarówno wobec naszych, jak i cudzych preferencji. Z tego względu wśród materialnych elementów istotnych dla trzeciego etapu myślenia krytycznego należy wyróżnić preferencje wszystkich osób zaangażowanych w sytuację etyczną oraz ich siłę⁶⁰.

Powyższy warunek pełnego utożsamienia się z innym, wraz z nieistotnością stałych indywidualnych, powoduje ważne normatywne konsekwencje. Jeśli dokonujemy moralnego namysłu np. nad takimi problemami, jak kontrola urodzeń⁶¹, antykoncepcja, aborcja⁶², eksperymenty na embrionach⁶³ czy zapłodnienie *in vitro*⁶⁴, „gdy mamy wybór między spowodowaniem, aby ktoś zaistniał, a nieuczynieniem tego, interes tej możliwej osoby musi być wzięty pod uwagę”⁶⁵. Uniwersalizacja zmusza nas bowiem, abyśmy rozważyli hipotetyczne sytuacje, a w konsekwencji preferencje możliwych osób, nawet tych, które aktualnie nie istnieją⁶⁶, oraz traktowali je na na równi z interesami osób żyjących (przy uwzględnieniu prawdopodobieństwa ich zaistnienia, co jest ważne dla wyznaczenia zakresu bliskich świa-

⁶⁰ Zob. tamże, s. 15.

⁶¹ Zob. tenże, *Possible People*, w: tenże, *Essays on Bioethics*, dz. cyt., s. 67-83.

⁶² Zob. tenże, *Abortion and the Golden Rule*, w: tenże, *Essays on Bioethics*, dz. cyt., s. 147-167; tenże, *A Kantian Approach to Abortion*, w: tenże, *Essays on Bioethics*, dz. cyt., s. 168-184.

⁶³ Zob. tenże, *Embryo Experimentation*, w: tenże, *Essays on Bioethics*, dz. cyt., s. 118-130.

⁶⁴ Zob. tenże, *In Vitro Fertilization*, w: tenże, *Essays on Bioethics*, dz. cyt., s. 98-117.

⁶⁵ Zob. tenże, *Possible...*, dz. cyt., s. 67.

⁶⁶ Krytykę tezy, iż nieistniejące, jedynie możliwe osoby są podmiotem myślenia krytycznego (jako formy tzw. złotej reguły), można znaleźć w: G. Sher, *Hare, Abortion, and the Golden Rule*, „Philosophy and Public Affairs” (2) 1977, s. 185-190.

tów możliwych, a co za tym idzie, sytuacji hipotetycznych, które są istotne). Podczas krytycznego namysłu nad aborcją do moralnie istotnych należą np. preferencje, które możemy poznać, jeśli zadamy sobie pytanie, czy chcielibyśmy, aby nasza matka dokonała aborcji na naszej osobie. Gdy zastanawiamy się nad zapłodnieniem *in vitro*, między innymi powinniśmy postawić sobie pytanie, czy jeśli przyszlibyśmy na świat dzięki takiej metodzie, pragnęlibyśmy, aby ów zabieg nie został był dokonany. Ponieważ zdaniem Hare’a większość ludzi, którzy egzystują powyżej minimalnego poziomu zadowolenia z życia, na powyższe pytania odpowiedziałoby negatywnie, takie praktyki, jak antykoncepcja, aborcja, celibat czy eksperymenty na embrionach, nie są moralnie obojętne. Choć ich ocena jest kwestią o wiele bardziej skomplikowaną niż odpowiedź jedynie na powyższe pytania, to generalnie rzecz ujmując, jeśli nie weźmie się pod uwagę płynących z tych praktyk korzyści oraz pewnych dodatkowych związanych z nimi faktów, stwierdzić należy, że każda praktyka niszcząca potencjalne życie istot, które mają zdolność odczuwania, jest „wewnętrznie moralnie zła”⁶⁷. Powyższa wstępna ocena nie wynika jednak z nadawania szczególnej metafizycznej wartości ludzkim komórkom rozrodczym czy z przekonania, że płód to również istota ludzka, lecz z zaprzepaszczenia szansy na ich rozwinięcie się w istotę, która będzie posiadać spełnialne preferencje⁶⁸. Owo „wstępne zło” jest tym większe, im większa byłaby owa szansa.

Jeśli przez formalność rozumie się analityczność, myślenie krytyczne nie jest czysto formalne – poza pojęciowymi prawami języka moralności powiązane jest z założeniami dodatkowymi. Oprócz tezy metaetycznych zakłada ono: racjonalność podmiotów działających jako racjonalność instrumentalną oraz *warunek pełnego utożsamienia się z innym*⁶⁹. Należy posiadać również wiedzę na temat istotnych materialnych własności (*mRNat*) dotyczących ocenianej

⁶⁷ Zob. R. M. Hare, *When does Potentiality Counts?*, w: tenże, *Essays on Bioethics*, dz. cyt., s. 91.

⁶⁸ Traktowanie interesów potencjalnych istot rodzi pewne teoretyczne trudności (np. postulowany przez krytyków obowiązek wzrostu populacji), z których Hare zdawał sobie sprawę i które rozważał, w: tenże, *Possible...*, dz. cyt.; tenże, *When does Potentiality...*, dz. cyt.

⁶⁹ A. Gibbard warunek ten nazywa *Conditional Reflection Principle*. Zob. A. Gibbard, *Hare's Analysis of 'Ought'*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 60.

sytuacji, aby tej metody użyć do oceny konkretnego przypadku⁷⁰. Gdy jednak za Hare'em przyjmiemy się różne od analityczności czy tautologiczności, luźniejsze znaczenie słowa formalność (niesuponowanie żadnych konkretnych subiektywnych czy obiektywnych treści materialnych – tego, co ludzie preferują i czego pragną), można argumentować, że opisane w § 5.2, odpowiedzialne za trzeci etap myślenia krytycznego, prawo superwencji HSup*II:

$$\text{HSup*II} \quad \begin{aligned} & \Box((\exists x)(\text{Mor}_x \& \text{RNat}_x \& (\text{RNat}_x \text{ Sup Mor}_x)) \Box (y) (\text{RNat}_y \Box \text{Mor}_y)) \\ & \sim \Box(\exists x) (\text{RNat}_x \& \sim \text{Mor}_x) \end{aligned}$$

również jest regułą formalną. Hare, tak jak i Kant, zastanawiał się bowiem,

jakie maksymy, czyli podstawowe zasady, mogłyby przyjąć zbiorowość podmiotów, aby zarazem nie zakładać czegokolwiek na temat pragnień podmiotów działających i stosunków społecznych⁷¹.

Z tego względu metodę tworzenia zasad moralnych można nazwać formalną.

Zgodnie z tym, co pisał Hare, z trzema etapami myślenia krytycznego związane są dwa aspekty utilitaryzmu, ufundowane na dwu rodzajach bezstronności. Przez **formalny aspekt utilitaryzmu** Hare rozumiał te własności swojej etyki, które maksymalnie zbliżają ją do Kanta. Formalny aspekt utilitaryzmu jest przeformulowaniem żądania, aby nasze preskrypcje, które tworzą zasady moralne, były uniwersalne formalnie⁷². **Utilitaryzm materialny** odpowiada natomiast wszystkim tym aspektom jego etyki normatywnej, które nie mają analitycznego charakteru i zostały założone na podstawie obserwacji świata aktualnego. Z powyższym podziałem wiążą się dwa rodzaje bezstronności. Formalna zasada superwencji wraz z preskryptywnością w wyniku pierwszego i drugiego etapu uniwersalizacji własności $fRNat$ prowadzi do **bezstronności formalnej**, natomiast przy uwzględnieniu cech $mRNat$ (trzeci etap uniwersalizacji) do **bezstronności materialnej**⁷³. Jednak jak już wskazaliśmy, analityczność drugiego etapu myślenia moralnego może zostać podważona, gdyż w tym kroku Hare zakładał racjonalność podmiotów moralnych.

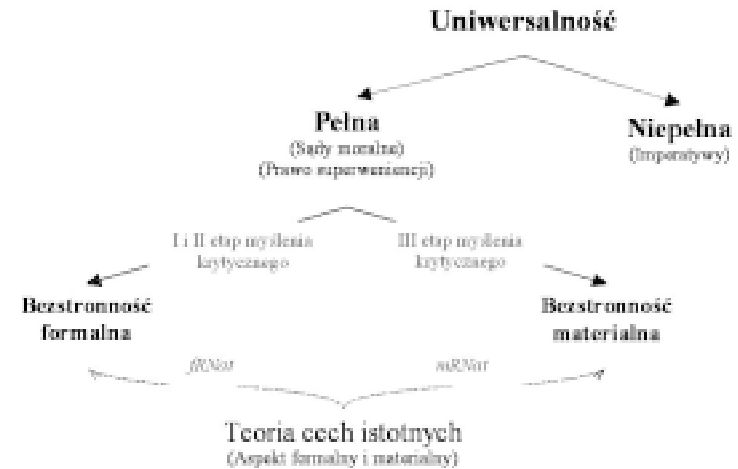
⁷⁰ Zob. R. M. Hare, *Comments on Nagel...*, dz. cyt., s. 248-249.

⁷¹ O. O'Neill, *Etyka Kantowska*, dz. cyt., s. 214.

⁷² Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 15-16, 68.

⁷³ Zob. tenże, *Comments on Nagel...*, dz. cyt., s. 249-250.

Relację między różnymi etapami myślenia krytycznego, bezstronnością, uniwersalnością i dwoma aspektami utilitaryzmu można zobrazować w następujący sposób:



Rysunek 9. Rodzaje uniwersalności wg R. M. Hare'a

§ 6.3.4 Racjonalność

Idealna metoda myślenia moralnego wymaga od nas umiejętności posługiwania się rozumem w ramach ograniczeń związanych z formalnymi cechami moralnej gry językowej. Dyskurs moralny, racjonalność moralna to bowiem racjonalność praktyczna, rozważa ograniczona przez preskryptywność, uniwersalność oraz nadrzędność, tj. cechy logiki moralności. Racjonalizm koncepcji Hare'a objawia się również tym, iż dziedzina moralnie relewantnych preferencji ograniczona jest ze względu na ich racjonalność. Zdaniem Hare'a, moralnie istotne są bowiem jedynie preferencje, które są racjonalne.

Zanim uzasadnimy powyższe stwierdzenie, wyjaśnić jednak należy, co Hare rozumiał przez racjonalność. Niejednokrotnie używał on słowa „rozważa”, „roztropność” zamiennie ze słowem „racjonal-

ność”. Przez **rozwagę** lub **roztropność** rozumiemy pojęcie, któremu odpowiada angielski termin *prudence*.

Być rozważnym to tyle co o przyszłych stanach pewnej osoby (zazwyczaj osoby, której ciało stanie się naszym obecnym ciałem) myśleć jak o własnych, a tym samym budować troskę o działania zgodne z przyszłymi preferencjami tej osoby. Być moralnym to przede wszystkim rozważyć hipotetyczną sytuację, w której to, co rzeczywiście będzie stanami innej osoby, stałoby się naszymi stanami, a tym samym budować hipotetyczną troskę o spełnienie własnych preferencji w tej hipotetycznej sytuacji; a następnie, ze względu na uniwersalizację, poczuć, że jest się zmuszonym (chyba że wybiera się drogę ucieczki amoralisty) do przeobrażenia tej wyłącznie hipotetycznej troski w troskę rzeczywistą, gwołi spełnienia preferencji rzeczywistej drugiej osoby⁷⁴.

Chcąc zatem wydać moralną preskrypcję, powinniśmy wziąć pod uwagę nie to, co my i inni aktualnie preferujemy i czego pragniemy, ale czego powinniśmy pragnąć, jeśli mielibyśmy być roztropni. Zdaniem Hare’a, **osądzić racjonalnie** to osądzić na podstawie pełnych informacji, nie popełniając błędów logicznych. **Racjonalny wybór** danej osoby to taka najkorzystniejsza dla niej decyzja, którą podjąłby każdy, kto wziąłby pod uwagę wszystkie potrzebne, relewantne informacje i byłby na miejscu tej osoby⁷⁵. **Racjonalna preferencja** danej osoby to preferencja, którą posiadałby każdy, kto byłby na jej miejscu i miał pełną wiedzę i niezmacony umysł⁷⁶.

Na pytanie, jaki byłby przykład konkretnej racjonalnej preferencji, nie istnieje uniwersalna odpowiedź, używane bowiem przez Hare’a pojęcie racjonalności jest beztreściowe. Nie znaczy to jednak, że w praktyce jest ono puste. Pokrywa się z pojęciem **racjonalności instrumentalnej**, która polega na optymalnym doborze środków do wyznaczonych celów, przy czym cele owe nie podlegają już racjonalnej selekcji. Z tego względu pojęcie racjonalnej preferencji czy decyzji zapełnia się treścią dopiero po wzięciu pod uwagę konkretnej sytuacji danej osoby i żywionych przez nią pragnień – celów.

Istnieje jednak kilka rodzajów preferencji wyróżnionych ze względu na odniesienie czasowe, pojawia się więc pytanie, które z nich są istotne dla racjonalnego jak i moralnego namysłu. Wskazać można np. na preferencje synchroniczne i asynchroniczne. Pierwsze z nich

⁷⁴ R. M. Hare, *Myślenie...*, dz. cyt., s. 273.

⁷⁵ Powyższy sposób pojmowania racjonalności Allan Gibbard nazwał *Full Information Analyses* i poddał krytyce w: A. Gibbard, *Wise Choices...*, dz. cyt., s. 18-22.

⁷⁶ Zob. R. M. Hare, *Ethical Theory and Utilitarianism*, dz. cyt., s. 218.

odnoszą się do pragnień i upodobań występujących w danym czasie *t*, aby w tym właśnie czasie zostać zaspokojone, drugie zaś do odmiennych punktów czasowych. Ową złożoność rodzajów preferencji przedstawia poniższa tabela 6.

preferencje synchroniczne	1	teraz-na-teraz	aktualna preferencja dotycząca teraźniejszości (chciałbym teraz, aby przestała boleć mnie głowa)
	2	wtedy-na-wtedy	przeszła preferencja dotycząca przeszłości (kiedyś chciałem, aby w tamtym czasie przestała boleć mnie głowa)
	3	kiedyś-na-kiedyś	przyszła preferencja dotycząca przyszłości (w czasie, gdy będę stary, będę chciał przejść na emeryturę)
preferencje asynchroniczne	4	teraz-na-kiedyś	aktualna preferencja dotycząca przyszłości (chciałbym dziś, abym w wieku sześćdziesięciu lat przeszedł na emeryturę)
	5	teraz-na-wtedy	aktualna preferencja dotycząca przeszłości (żałuję, że nie uczyłem się kiedyś fizyki; dziś chciałbym, abym się bardziej przykładał był do fizyki w przeszłości)
	6	wtedy-na-teraz	przeszła preferencja dotycząca sytuacji teraźniejszej (kiedyś chciałem mieć 18 lat i dziś mam osiemnaste urodziny)
	7	wtedy-na-kiedyś	przeszła preferencja, dotyczące sytuacji przyszłej (kiedyś chciałem umrzeć w wieku 100 lat)
	8	kiedyś-na-teraz	przyszła preferencja dotycząca teraźniejszości (kiedyś będę zakował tego, co chcę dzisiaj; kiedyś będę chciał, abym w dniu aktualnym zachował się inaczej)
	9	kiedyś-na-wtedy	przyszła preferencja dotycząca przeszłości (kiedyś będę chciał, abym, gdy byłem młody, przykładał był się bardziej do fizyki)

Tabela 6. Rodzaje preferencji

W związku z powyższym podziałem pojawiają się pytania dodatkowe: W jaki sposób powinniśmy ważyć wymienione rodzaje preferencji między sobą? Jak porównywać nasze aktualne preferencje z przyszłymi, oraz aktualne preferencje z alternatywnymi przyszłymi preferencjami, jakie będziemy mieli, jeśli dokonamy alternatywnego wyboru? Którą z dostępnych opcji działania powinniśmy aprobować i według jakiej metody wybrać opcję najlepszą?

Według Hare'a, dokonując racjonalnego wyboru, istotne jest brać pod uwagę nie tylko terażniejszych preferencji. Założmy, że jakieś dziecko aktualnie pragnie zjeść kolejne pudełko słodczy, jednakże spowoduje to jego problemy z nadwagą, czego później będzie żałował. W takich przypadkach należałoby zadać pytanie, które działanie promowałoby zaspokojenie wszystkich jego preferencji⁷⁷. Człowiek roztropny nie przedkłada bowiem aktualnych preferencji teraz-na-teraz nad preferencje przyszłe teraz-na-kiedyś i kiedyś-na-kiedyś. Wszystkie możliwe i aktualne preferencje traktuje jako równoważne, starając się je zmaksymalizować, bacznie jedynie na prawdopodobieństwo ich spełnienia oraz ich siłę. Osoba taka stara się również brać pod uwagę faktyczne przyszłe preferencje (kiedyś-na-kiedyś), a nie te, o których dziś się jej wydaje, że w przyszłości będą jej udziałem (teraz-na-kiedyś).

Dla archanioła preferencje teraz-na-kiedyś i kiedyś-na-kiedyś zawsze posiadać będą identyczną treść i siłę, nie myli się on bowiem co do przyszłości ani nie zmienia zdania. Człowiek, podejmując decyzje moralne, domyśla się jedynie swoich przyszłych preferencji, niejednokrotnie reorganizuje i na nowo wyznacza cele, na przestrzeni czasu ulega zmianom. Ludzkie aktualne pragnienia dotyczące przyszłości często nie pokrywają się z przyszłymi pragnieniami dotyczącymi przyszłej terażniejszości, dlatego też trudno jest nam szacować całościowe spełnienie naszych preferencji. Osoba racjonalna stara się jednak formułować takie preferencje teraz-na-kiedyś, które pokrywają się z przyszłymi preferencjami kiedyś-na-kiedyś. Zdaniem Hare'a, przykładem racjonalnego wyboru jest decyzja Odyseusza, który spodziewając się słyszeć śpiew Syren i przewidując swoją na ów śpiew reakcję, zażądał, aby go związano. Jego wybór był racjonalny, ponieważ gdyby zrealizował swoją przyszłą bardzo silną preferencję, aby udać się za głosem Syren, rozbiłby swój statek i nie spełnił innych silnych preferencji, które żywi teraz i które będzie miał później, np. aby żyć, zobaczyć Itakę i Penelopę. Według Hare'a, Odyseusz podjął racjonalną decyzję na podstawie całkowitej siły swoich preferencji⁷⁸. Choć Hare miał pewne wątpliwości, ostatecz-

⁷⁷ Zob. tenże, *An Utilitarian Approach to Ethics*, w: tenże, *Objective prescriptions...*, dz. cyt., s. 157.

⁷⁸ Zob. tenże, *Some Reasoning about Preferences*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 247.

nie skłaniał się do przekonania, iż podczas krytycznego namysłu moralnego nie są istotne preferencje asynchroniczne. Według niego, jeśli utylitaryzm szczęścia, którego adwokatem był R. Brandt, przeformułujemy jako jeden z rodzajów utylitaryzmu preferencji, tzn. jako teorię, wedle której powinno się maksymalizować spełnienie preferencji synchronicznych, można powiedzieć, że taka teoria jest najbardziej adekwatną koncepcją normatywną⁷⁹.

Powyższe zasady dotyczą racjonalnego wyboru o charakterze egoistycznym. Jednak podczas namysłu moralnego musimy brać pod uwagę również preferencje innych osób. J. L. Mackie w jednej ze swoich uwag krytycznych do etyki Hare'a twierdził, że już na drugim, nie mówiąc o trzecim etapie uniwersalizacji, wątpliwe jest, czy istnieją zasady, które przeszłyby wymagania Hare'a: sytuacje ludzi są bowiem bardzo różne i nie sposób zadowolić wszystkich⁸⁰. Identyfikacyjny sposób argumentacji przedstawia Gilbert Harman, twierdząc, iż nie istnieje racjonalny zbiór zasad, który byłby akceptowany przez każdego⁸¹. Hare nie szukał jednak reguł moralnych, które nie naruszyłyby niczyich interesów i spełniały preferencje każdej osoby. Według niego, aby nasze decyzje moralne były racjonalne oraz zgodne z logiką pojęć, musimy wziąć pod uwagę preferencje wszystkich osób zaangażowanych w sytuację etyczną, a w razie konfliktu, bezstronnie wybrać z nich najsilniejszą. Czyn, który zaspokoi taką najsilniejszą preferencję, będzie tym, który z perspektywy moralności powinniśmy wybrać⁸². Powyższe zalecenie uzasadniał on opierając się na założeniach dotyczących racjonalności indywidualnej. Gdy stajemy przed podobnym problemem, jeśli w grę wchodzi nasze własne sprzeczne preferencje, wybieramy realizację tej z nich, która jest silniejsza. Jego zdaniem, taka decyzja z perspektywy indywidualnej często jest racjonalna, dlatego aby rozwiązać jakiś problem praktyczny, czy o charakterze moralnym, czy indywidualnym, należy porównać ze sobą siłę konkurencyjnych aktów woli, a następnie wybrać takie działanie, które zrealizuje najsilniejszy z nich⁸³.

⁷⁹ Zob. tenże, *Preferences of Possible People*, w: tenże, *Objective Prescriptions...*, dz. cyt., s. 126-127; tenże, *A Utilitarian Approach...*, dz. cyt., s. 158-159.

⁸⁰ Zob. J. L. Mackie, *Ethics...*, dz. cyt., s. 93.

⁸¹ Zob. G. Harman, *The Nature of Morality. An Introduction to Ethics*, Oxford University Press, New York 1977, s. 81.

⁸² Zob. R. M. Hare, *Relevance...*, dz. cyt., s. 210-211.

⁸³ Zob. tenże, *Some Reasoning...*, dz. cyt., s. 249.

Zalecenie, aby wybrać najsilniejszą preferencję, jest racjonalne jedynie wtedy, gdy dotyczy sytuacji, w których każda z nich jest w nierozstrzygalnym konflikcie, tak że tylko jedna może być spełniona i spełnienie ich nie powoduje dalszych istotnych następstw. Zazwyczaj istnieją jednak opcje działania, które zaspokajają kilka preferencji na raz i powodują istotne odległe skutki. Załóżmy, że mamy dwie możliwości: opcję A i B. Załóżmy dalej, że w grę wchodzi trzy istotne preferencje, które musimy wziąć pod uwagę, oznaczone małymi literami *a*, *b* i *c*. Siła każdej z nich oscyluje w granicach od jeden do dziesięciu. Jeśli dostępne opcje działania przedstawiają się w następujący sposób:

Opcje	A		B
Siła preferencji	<i>a</i>	<i>b</i>	<i>c</i>
	5	4	7
Suma	9		7

Tabela 7. Przykładowe opcje wyboru

powinniśmy, zdaniem Hare'a, **maksymalizować** spełnienie preferencji, czyli wybrać opcję A. W większości sytuacji racjonalne działanie polega na spełnieniu jak największej ilości jak najsilniejszych preferencji, co jest zgodne z rozumieniem racjonalności przez klasyków utilitaryzmu. Jeśli racjonalność wyraża się w optymalnym doborze środków do wyznaczonych celów, przy czym celem owym jest spełnienie posiadanych preferencji, to optymalne jego zaspokojenie wydaje się być realizacją maksymalnej ich ilości. Przy tym podejściu preskrypcja: „Spełnij maksymalną ilość twoich aktualnych i przyszłych preferencji!” jest wolą rozsądku. Jeden z ważniejszych argumentów uzasadniających kryterium maksymalizacji użyteczności głosi, że koresponduje ono z racjonalnością praktyczną, gdyż to maksymalizacja pożądaných wartości jest najbardziej racjonalną na nie reakcją⁸⁴.

⁸⁴ Zob. T. Mulgan, *The Demands of Consequentialism*, Clarendon Press, Oxford 2001, s. 14.

Wybierając opcję B, urzeczywistnilibyśmy wariant, w którym zarówno ilość spełnionych preferencji, jak i suma ich sił są mniejsze, co byłoby nieroztropne.

Według Richarda Hare'a, tak jak racjonalne jest maksymalizowanie sum preferencji egoistycznych, tak też podczas krytycznego namysłu moralnego istotne jest sumowanie i maksymalizacja spełnienia preferencji różnych osób. Podczas racjonalnego wyboru egoistycznego powinniśmy zrealizować nasz interes, który odkrylibyśmy, jeśli wzięlibyśmy pod rozwagę nasze faktyczne preferencje synchroniczne oraz poddali je testowi logiki i faktów. Stosowanie myślenia krytycznego, a zatem bezstronne podchodzenie do stanów motywacyjnych różnych ludzi, nie zawsze jest jednak racjonalne ze względu na nasze egoistyczne interesy. Jak wyjaśniliśmy w rozdziale czwartym, można być racjonalnym amoralistą, sądy moralne nie są bowiem nadrzędne w tym sensie, że dostarczają kategorię racji, aby ich przestrzegać. Pragnąc dokonać moralnego wyboru na płaszczyźnie krytycznej, musimy skonstruować uniwersalną zasadę, która maksymalizowałaby spełnienie interesów wszystkich zaangażowanych osób. Hare był zdania, że procedura sumowania siły i ilości spełnionych preferencji zaangażowanych osób jest konsekwencją bezstronności, która jest ufundowana na uniwersalności sądów moralnych⁸⁵. Bezstronność w etyce objawia się równym traktowaniem interesów poszczególnych osób, co w praktyce oznacza przedkładanie zaspokojenia dwóch preferencji dwóch osób nad jedną preferencją jednej (przy założeniu, że każda z nich ma taką samą siłę). **Agregacjonizm**, cecha utilitaryzmu, która głosi że należy sumować spełnienie preferencji różnych osób, zdaniem Hare'a jest najbardziej racjonalną odpowiedzią na wymagania bezstronności stawiane przez logikę pojęć moralnych⁸⁶.

Z powyższymi cechami utilitaryzmu wiąże się pewien zarzut. Według Johna Rawlsa, utilitaryści popełniają błąd rozciągania zasad racjonalności indywidualnej na dziedzinę stosunków społecznych⁸⁷. Nawet jeśli zgodzimy się, że indywidualna racjonalna decyzja

⁸⁵ Zob. R. M. Hare, *Relevance...*, dz. cyt., s. 209; tenże, *Ethical Theory and Utilitarianism*, dz. cyt., s. 215.

⁸⁶ Zob. tenże, *An Utilitarian Approach...*, dz. cyt., s. 159-160.

⁸⁷ Zob. J. Rawls, *Teoria sprawiedliwości*, tłum. M. Panufnik, J. Pasek, A. Romanuk, PWN, Warszawa 1994, s. 45.

podlega powyżej zarysowanym zasadom, rodzi się wątpliwość, czy te reguły powinny mieć zastosowanie również przy wyborze moralnym. Przedmiotem indywidualnego namysłu egoistycznego są bowiem skonfliktowane preferencje, namysłu moralnego natomiast autonomiczne osoby. Dlaczego reguły wyboru między indywidualnymi preferencjami mają być tożsame z zasadami wyboru między interesami odrębnych osób? Argument ten może mieć interpretację metafizyczną: człowiek jest osobą, a preferencja stanem umysłu, dlatego nie należy przenosić procedury wyboru z preferencji na ludzi. Jednak zdaniem Hare'a, tak jak większości utilitarystów, jedynie preferencje czy interesy są moralnie istotne: struktura bytu, który je posiada, nie jest ważna⁸⁸. Namysł moralny nie polega na rozstrzygnięciu między metafizycznymi bytami, lecz ich interesami: różnice metafizyczne nie odgrywają w tym wypadku roli. Z tego względu można zaproponować również inne odczytanie powyższego zarzutu: chęć zaspokojenia interesów, która jest przedmiotem ludzkich dążeń, jest innym rodzajem motywacji, niż dążenie do spełniania preferencji, które nie leżą w naszym interesie. Utylityzm nie traktuje ludzi jako odrębnych jednostek, nie bierze bowiem pod uwagę ich interesów, które są ufundowane na ich preferencjach, lecz odrywa preferencje od osób, czyniąc je bezosobowym przedmiotem maksymalizacji.

Dla przykładu załóżmy, że musimy rozstrzygnąć problem moralny, w który zaangażowane są trzy osoby: Jan, Krzysztof i Małgorzata. Występujący interpersonalny konflikt preferencji rozwiązać możemy wyłącznie za pomocą trzech uniwersalnych zasad, które będą stanowić podstawę naszego wyboru. Oznaczmy je literami A, B i C. Aby uprościć schemat problemu, weźmiemy pod uwagę jedynie preferencje teraz-na-teraz i kiedyś-na-kiedyś, których siłę reprezentować będą liczby od 1 do 10. Ów konflikt preferencji ilustruje tabela 8, natomiast rozkład moralnie najkorzystniejszych opcji ze względu na sposób interpretacji słusznego wyboru tabela 9:

Opcjonalne zasady moralne	Jan			Krzysztof			Małgorzata											
	A	B	C	A	B	C	A	B	C									
Sila preferencji teraz-na-teraz	2	4	2	2	3	2	2	1	5	6	2	3	4	2	2	3	3	5
Sila preferencji kiedyś-na-kiedyś	3	1	1	1	7	4	1	7	3	3	2	3	3	3	3	1	3	2
Suma sil spełnionych preferencji	27	13	31	25	33	17	27	20	29									
Egoistycznie najkorzystniejsza zasada (indywidualny interes)			X			X												X

Tabela 8. Interpersonalny konflikt preferencji

Opcjonalne zasady	A	B	C
Opcja spełniająca najbliższą racjonalną preferencję zaangażowanej osoby		33 (Krzysztof)	
Suma spełnionych faktycznych preferencji Jana, Krzysztofa i Małgorzaty	79 (opcja małej maksymalnej liczby spełnionych faktycznych preferencji)	66	76
Suma spełnionych racjonalnych preferencji Jana, Krzysztofa i Małgorzaty	0	30	60 (opcja małej maksymalnej liczby spełnionych racjonalnych preferencji)

Tabela 9. Rozkład moralnej słuszości opcji ze względu na sposób interpretacji moralnie racjonalnej decyzji

Z tabeli 8 wynika, iż najkorzystniejszą zasadą, ze względu na interes Jana i Małgorzaty jest zasada C, dla Krzysztofa jest nią zaś B. Reprezentują one indywidualny interes poszczególnych osób. W świetle drugiej interpretacji powyższego zarzutu Rawlsa dla utilitarystów nie jest jednak istotne, czy promowana reguła pokrywa

⁸⁸ Zob. R. M. Hare, *Freedom...*, dz. cyt., s. 222.

się z egoistycznym interesem osób zaangażowanych – z ich racjonalnymi preferencjami. Według niego, w świetle utilitaryzmu najlepsza zasada musi być opcją maksymalizującą spełnienie wszystkich, bezstronnie ocenianych preferencji. Z tabeli 9 wynika, że zasadą maksymalizującą spełnienie faktycznych preferencji wszystkich osób jest reguła A. Nie pokrywa się ona jednak z interesem żadnej ze stron z tabeli 8. Jeśli zasady racjonalności wymagają maksymalizowania spełnienia preferencji, zasada B, choć wyraża bezstronnie najsilniejszy indywidualny interes (racjonalną preferencję Krzysztofa), musi zostać odrzucona.

Powyższa interpretacja zarzutu Rawlsa nie dotyka jednak teorii Hare'a. Według niego, jeśli dla oceny moralnej istotnymi cechami naturalnymi są preferencje, to, jak już wspomnieliśmy, tylko te z nich, które są racjonalne, tylko te, które nazwać można interesem.

Kiedy mówiłem [...], iż muszę uczynić dla każdego człowieka, który wchodzi w zakres mojego działania, to co chciałbym, aby było uczynione mi, mówiłem nieściśle. [...] Kiedy wydaję zalecenie w swoim własnym czy cudzym interesie [...] muszę zapytać, nie o to, czego aktualnie pragniemy, lecz o to, czego z perspektywy roztropności powinniśmy chcieć. Z takiego racjonalnego punktu widzenia (racjonalności jako roztropności) muszę wydać uniwersalne zalecenia. [...] Kiedy rozważam pragnienia innych, rozważam je, zakładając, jakie one by były, gdyby osoby te były doskonale rozważne, tzn. myślę o pragnieniach, które osoby te by posiadały, gdyby były w pełni poinformowane i nie miały zmażonego umysłu. [...] Musimy pamiętać, iż gdy będę mówił o „pragnieniach”, „zaleceniach” itd., będę miał na myśli „pragnienia”, „zalecenia” z punktu widzenia osoby, która jest rozważna na tyle, aby realizować swój interes⁸⁹.

Zdaniem Hare'a, reguły moralne powinny maksymalizować racjonalne ludzkie wybory, przyczyniać się do zwiększania ilości spełnionych racjonalnych pragnień. Z tego względu zasada moralna, którą powinniśmy przyjąć, ma postać reguły C, co przedstawia tabela 9. W przeciwnym razie maksymalizowalibyśmy preferencje, które nie są racjonalne, które nie leżą w interesie osób je posiadających.

Racjonalna preferencja (interes) osoby O jest to preferencja, którą posiadałby każdy, kto byłby na jej miejscu i miał pełną wiedzę i niezmażony umysł. Choć pojęcie to jest skonstruowane i ma charakter idealizacji, to jednak jego treść niekiedy pokrywa się z faktycznie istniejącymi preferencjami. W innych przypadkach taka preferencja może być nieświadoma i nieprzeżywana przez żadną osobę,

⁸⁹ R. M. Hare, *Ethical Theory and Utilitarianism*, dz. cyt., s. 217-218.

pomimo iż superwenuje na faktycznych, doświadczanych preferencjach danej osoby, np. w naszym interesie, biorąc pod uwagę niezwerbalizowane, odległe i silne preferencje kiedyś-na-kiedyś, może być dokonanie czynu, którego nie mamy ochoty dziś wykonać, a nawet o którym nie mamy świadomości, że może być spełniony. Pamiętać należy, że interes, o którym mówimy, zawsze musi bazować na subiektywnych stanach wolitywnych danej osoby. Hare, jako utilitarysta akceptujący welfaryzm⁹⁰, nie zgodziłby się z rozumieniem interesu jako dobra niezależnego od naszych faktycznych preferencji teraźniejszych i przyszłych. Według niego, nie można byłoby np. powiedzieć, że w moim interesie leży to, abym żył dłużej, pomimo iż moje silne obecne, tak samo jak bardzo silne przyszłe preferencje wskazują, że rozwiązaniem maksymalizującym spełnienie wszystkich moich preferencji jest dokonanie samobójstwa.

Jeśli podczas trzeciego etapu myślenia krytycznego relewantnymi cechami mają być pierwszoosobowe stany motywacyjne, takie jak pragnienia, uczucia, preferencje, chęci, rodzi się wątpliwość, czy teoria Hare'a nie sprowadza myślenia moralnego, wbrew jej autorowi, do porównywania siły idei i przekonań moralnych. Krytyka zastanych, domniemych przekonań moralnych w świetle procedury myślenia krytycznego ma bowiem uzasadnienie o tyle, o ile nasze preferencje i chęci są niezależne od przekonań. Nie wydaje się to jednak prawdą. Preferencje są powiązane w sposób istotny z żywionymi sądami. Jeśli mamy z całą powagą ważyć pragnienia muzułmanina, fanatycznego nazisty, postmodernistycznego nihilisty, chrześcijanina, satanisty, etycznego utilitarysty i obrońcy życia, może się okazać, że porównywać będziemy nie siłę preferencji, lecz siłę moralnych intuicji i wiary w wyznawany światopogląd. Intuicje etyczne, od których chciał abstrahować Hare, istotnie wpływają bowiem na nasze preferencje. Można więc twierdzić, że jeśli nie uda się wyznaczyć linii demarkacyjnej między „czystymi” preferencjami a preferencjami związanymi z wartościami i przekonaniem etycznymi czy religijnymi i abstrahować od tych drugich, to w konsekwencji Hare, choć bardzo się od tego wzbraniał, jest skazany na intuicjonizm. Z powyższą wątpliwością powiązany jest kolejny zarzut, który postawił H. M. Robinson⁹¹. Według niego,

⁹⁰ Opis welfaryzmu znajduje się w § 6.5.

⁹¹ Zob. H. M. Robinson, *Is Hare a Naturalist?*, „The Philosophical Review” (1) 1982, s. 80-81.

traktowanie wyznawanych wartości i przekonań innych osób na równi z naszymi nie jest racjonalne, jeśli jesteśmy przeświadczeni, że ich punkt widzenia jest fałszywy. Jeśli posiadane preferencje zależne są od akceptowanych idei, wyznawane idee są natomiast zbiorem przekonań, czemu mielibyśmy traktować na równi przekonania fałszywe i prawdziwe, a rozstrzygnięć dokonywać na podstawie ich siły czy maksymalizacji ich spełnienia?

Odpowiadając na powyższe zarzuty, można bronić tezy, że archaniot, aby rozstrzygnąć skonfliktowane preferencje np. teisty i satanisty, nie musiałby polegać na wazeniu intuicji. Jeśli bowiem istotne są jedynie preferencje racjonalne i założymy, że różnice światopoglądowe i akceptowane ideały wpływające na odmienne pragnienia w przytłaczającej części są ufundowane na różnicach w przekonaniach, których prawdziwość lub fałszywość idealny podmiot mógłby rozstrzygnąć, to poddając sprzeczne preferencje teisty i satanisty testowi logiki i faktów, mogłoby się okazać, że tylko pragnienia jednego z nich (albo żadnego) są racjonalne, a zatem istotne. Wydaje się, że na drodze eliminacji preferencji nieracjonalnych – preferencji, które nie leżą w interesie danej osoby – archaniot mógłby rozstrzygnąć dużą część trudnych do rozwiązania dla ludzi konfliktów. Z drugiej jednak strony, zasadnie można założyć, że nie wszystkie nasze preferencje są nieracjonalne z powodu nielogiczności czy bazowania na fałszywych przekonaniach. Niektóre nie są związane z weryfikowalnymi przekonaniem. Często pragniemy osiągnąć pewne cele dla nich samych i nie może podważyć ich konfrontacja z faktami. Nie wszystkie pragnienia mają instrumentalny charakter, służą realizacji danych celów, nie wszystkie podlegają zatem racjonalnej krytyce. Co więcej, według niektórych filozofów, np. I. Berlina czy B. Williama⁹², istnieją również preferencje, które tworzą wzory autentycznych, wewnętrznie skonfliktowanych wartości. Przy takich założeniach wybory musiałyby mieć charakter dramatycznego albo albo lub pozaracjonalnego zdeterminowania. Zdaniem Hare'a, jeśli archaniot musiałby jednak rozstrzygnąć konflikt między fundamentalnie różnymi, racjonalnymi czy pozaracjonalnymi, preferencjami-wartościami, powinien wybrać te, które są najsilniejsze. Głoszenie

⁹² Zob. B. Williams, *Konflikt wartości*, w: tenże, *Ile wolności powinna mieć wola i inne eseje z filozofii moralnej*, tłum. T. Baszniak i in., Aletheia, Warszawa 1999, s. 115-128.

nierozstrzygalnej dylematyczności niektórych konfliktów wartości jest rodzajem nadmiernego sceptycyzmu, który w świetle funkcji etyki jako praktycznego drogowskazu jest nie do przyjęcia. Podobnie niewspółmierne i niespójne są egoistyczne pragnienia, które rozwiązujemy, opierając się przy wyborze na całkowitej sumie sił preferencji.

W sytuacji konfliktu wartości czy idei istnieje jeszcze drugie rozwiązanie, z którego Hare zdawał sobie sprawę, lecz nie poświęcił mu wiele uwagi. Jak pisze, „z powodzeniem można też dążyć do zmiany nastawień społecznych, o ile miałyby to być zmiany najkorzystniejsza”⁹³. Preferencje ulegać mogą bowiem przeobrażeniom ze względu na wyznawane przekonania czy też techniki manipulacji. Jeśli racjonalne preferencje superwenują na preferencjach faktycznych, również one podlegają zmianom. Hare nie wyjaśnił jednak, co kryje się pod pojęciem najkorzystniejszej zmiany preferencji. Zastanawiając się, czy powinniśmy wpłynąć na interes danej osoby poprzez manipulację i zmianę jej **preferencji faktycznych** (faktycznych synchronicznych preferencji przeszłych, teraźniejszych i przyszłych), możemy odnieść się do jej idealnego interesu. Przez **idealny interes danej osoby** będziemy rozumieć własność, która jest ufundowana na maksymalizacji spełnienia jej idealnych preferencji. **Idealne preferencje** osoby O natomiast to preferencje, które przy wzięciu pod uwagę jej cech oraz cech świata są w najwyższym stopniu zoptymalizowane do bycia spełnionymi. Jeśli interes (racjonalna preferencja) jest pochodną preferencji faktycznych, to analogicznie, idealny interes jest pochodną idealnych preferencji (maksymalnie spełnialnych). Aby się dowiedzieć, czy zmiana preferencji jest korzystna, należy postawić pytanie, czy spowoduje ona wzrost sumy zaspokojonych nowych preferencji – czy zbliża osoby do posiadania preferencji idealnych. Zmiana jest korzystna, jeśli koszt dokonania zmiany, wyrażony ilością niezaspokojonych preferencji, będzie mniejszy niż różnica między ilością spełnionych nowych i preferencji starych.

Czym innym jest jednak płynąca ze zmiany preferencji korzyść indywidualna, a czym innym korzyść moralna. Czy jeśli instytucja niewolnictwa maksymalizowałoby faktyczne interesy wszystkich członków społeczeństwa, to z perspektywy moralnej należałoby ją

⁹³ R. M. Hare, *Myślenie...*, dz. cyt., s. 199-200; tenże, *Comments on Singer*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 269.

dla tego społeczeństwa zalecać, czy też, jeśli istniałby korzystniejszy dla wszystkich model społeczny, wymagający jednak istnienia odmiennych preferencji jej członków, powinniśmy dopasować instytucje do idealnych preferencji, które dziś nie są przeżywane? Jeśli korzystna zmiana w aspekcie indywidualnym jest liczona różnicą między zaspokojonymi przyszłymi preferencjami a preferencjami teraźniejszymi przy wzięciu pod uwagę kosztów takiej zmiany, to korzystna zmiana w aspekcie moralnym powinna zależeć od kosztów oraz różnicy między ilością i siłą zaspokojonych przyszłych interesów i interesów faktycznych wszystkich zaangażowanych osób? Zmiananie preferencji przebiega jednak bardzo wolno, często mając charakter ponadpokoleniowy. Pojawia się kolejne pytanie: jak ocenić zmianę, której koszty będzie ponosić jedno pokolenie, podczas gdy zyski kilka kolejnych? Jak długi horyzont czasowy, wylizując taki bilans, powinniśmy brać pod uwagę? Jeśli, jak twierdzi Hare'a, uniwersalizacja wymaga od nas postawienia się również w sytuacji potencjalnych osób, jako krytycznie myślący archaniołowie powinniśmy wziąć pod uwagę również interesy wszystkich nienarodzonych jeszcze ludzi. Archaniot, który miałby wiedzę o ilości takich osób, powinien za pomocą najkorzystniejszych metod dążyć do zaszczepienia w ludziach na przestrzeni lat takich preferencji, które będąc podstawą ich interesów, maksymalizowałyby spełnienie tych interesów przez wszystkie możliwe osoby. Archaniot utilitarysta powinien zatem, uwzględniając kondycję ludzką, naszą psychikę i fizjologię, dążyć do systematycznej, minimalizującej koszty zmiany naszych preferencji, tak aby w kolejnych pokoleniach zbliżały się one do idealnego zbioru interesów, które w optymalny sposób mogłyby z sobą współgrać. Taki stan, w którym wszyscy posiadają idealne interesy nazwać można **utilitarystycznym światem idealnym**. Zrealizowałyby się on wtedy, gdybyśmy posiadali idealne preferencje wyznaczone przez idealne interesy, których treść zależałaby od tego, czy w optymalny sposób maksymalizują spełnienie idealnych interesów wszystkich ludzi. Ponieważ zasady moralne muszą być nadrzędne nad zasadami egoistycznego sukcesu, w moralnie idealnym świecie linie fundowania i wyboru preferencji przebiegałyby daleko od idealnych zasad, przez idealne interesy do idealnych preferencji. Świat taki byłby utilitarystycznie najlepszym z możliwych światów.

Zarysowane powyżej zależności można zobrazować w następujący sposób:



Rysunek 10. Zmiana preferencji na podstawie utilitarystycznego świata idealnego.

Przedstawione powyżej zagadnienia dotyczące zmiany preferencji mają charakter wyłącznie sygnalizacyjny i wymagałyby odrębnych studiów. Niektóre, np. kwestia idealnego utilitarystycznego świata, nie zostały przez Hare'a w ogóle postawione i są związane z wieloma znakami zapytania. Nie używał on również pojęć idealnych preferencji czy idealnego interesu. Hare był bowiem przekonany, że problemy dotyczące racjonalnego namysłu wykraczają poza filozofię moralności. Z tego względu zadanie zanalizowane i bardziej szczegółowego opisanie tych zagadnień pozostawił innym. Jak napisał w tekście z 1998 roku w kontekście rozważanych problemów związanych z racjonalnością, „musimy teraz pozostawić te kwestie, w stosunku do których nie zająłem jeszcze stanowiska”⁹⁴. Nie przywiązywał do powyższego braku zbyt wielkiej wagi, dlatego że według niego, człowiek w swoim codziennym życiu, choć często nie potrafi podać procedury racjonalnego wyboru, umie podejmować racjonalne decyzje⁹⁵. Jeśli potrafimy w racjonalny sposób rozwiązywać codzienne wewnętrzne konflikty

⁹⁴ Tenże, *A Utilitarian Approach...*, dz. cyt., s. 159.

⁹⁵ Zob. tenże, *Ethical Theory and Utilitarianism*, dz. cyt., s. 217.

preferencji (pójść do kina czy poczytać książkę, spotkać się z przyjacielem czy popracować itd.), to nie widać przeszkód, abyśmy mogli skutecznie myśleć krytycznie. Z drugiej jednak strony, Hare zdawał sobie sprawę, że od kształtu teorii racjonalnej decyzji, która opisze naszą umiejętność racjonalnego rozwiązywania konfliktów, zależy konkretny kształt jego teorii normatywnej.

§ 6.4 Myślenie krytyczne a tożsamość osobowa

Jeśli jako archaniołowie przeprowadzimy trzy etapy uniwersalizacji, to utożsamimy się z preferencjami innych osób tak, iż staną się one naszymi własnymi. Skuteczne i zgodne z *warunkiem pełnego utożsamienia się z innym* przeprowadzenie trzeciego etapu myślenia krytycznego spowoduje, że naszymi stanami motywacyjnymi będą stany motywacyjne wszystkich osób zaangażowanych w sytuację etyczną. Nie istnieje wówczas konflikt pomiędzy preferencjami różnych osób (konflikt interpersonalny), ale niezgodność preferencji, które są naszymi własnymi (konflikt intrapersonalny)⁹⁶. Konsekwentne postawienie się na miejscu innej osoby, zgodnie z warunkiem pełnego utożsamienia się z innym, pociąga za sobą przyjęcie jej preferencji i motywacji jako naszych własnych⁹⁷. Z wymogiem tym związany jest jednak pewien teoretyczny problem. Czy jest metafizycznie możliwe pełne przeprowadzenie trzeciego etapu myślenia krytycznego – wzięcie pod uwagę subiektywnych stanów mentalnych wszystkich osób zaangażowanych w sytuację etyczną (ich preferencji, interesów, wierzeń, opinii itp.) – w taki sposób, aby uczynić je własnymi, bez utraty naszej integralności⁹⁸? Wydaje się, że nikt – nawet archanioł – nie może w pełni znaleźć się w „skórze innych”, nie gubiąc swojej tożsamości jako konkretnego indywiduum. Aby porównać naszą sytuację z sytuacją innej osoby musielibyśmy nią być, a to

⁹⁶ Zob. tenże, *Myślenie...*, dz. cyt., s. 140-141. Krytykę sprowadzanie konfliktów interpersonalnych do wewnętrznego konfliktu preferencji można znaleźć w: J. Griffin, *Interpersonal...*, dz. cyt., s. 71-82.

⁹⁷ Zob. R. M. Hare, *Relevance...*, dz. cyt., s. 210.

⁹⁸ Zob. C.C. W. Taylor, *Critical Notice on Freedom and Reason*, „Mind” (294) 1965, s. 280-298; H. S. Silverstein, *A Note on Hare on Imaging Oneself in the lace of Others*, „Mind” (323) 1972, s. 448-450. Odpowiedź na krytykę zawartą w powyższych artykułach znaleźć można w: M. H. Robins, *Hare's Golden-Rule Argument: A Reply to Silverstein*, „Mind” (332) 1974, s. 578-581.

jest niemożliwe. Nikt inny oprócz nas nie może patrzeć na świat z naszego punktu widzenia; nie można patrzeć z dwóch subiektywnych punktów widzenia. Jeśli założymy, że wiedza pierwszoosobowa, fenomenalne aspekty świadomości są nieredukowalne do opisu trzecioosobowego⁹⁹, wiedza na temat preferencji innych osób jest niemożliwa¹⁰⁰. Przy rozwiązywaniu konfliktów interpersonalnych Hare odwoływał się do ludzkiej umiejętności rozstrzygania konfliktów wewnętrznych¹⁰¹. Czy możemy jednak problemy międzyludzkie rozwikłać za pomocą reguł, jakie stosujemy do rozstrzygania wewnętrznych konfliktów preferencji, skoro preferencje innych nigdy w pełni nie staną się naszymi¹⁰²?

Niektórzy autorzy próbowali zatem wykazać, że metoda myślenia krytycznego, wyrażona metaforą „postawić siebie na miejscu innych”, jest metafizycznie niemożliwa¹⁰³. Byłoby to prawdą, gdybyśmy przyjęli którąś z rozpowszechnionych wersji deskryptywistycznej teorii identyczności osobowej. Przez taki pogląd rozumiemy, iż *Ja* sprowadzić można do opisu pewnych cech, bez względu na użyty do tego język: pierwszo lub trzecioosobowy, psychologiczny lub fizyczny. Najczęściej podawanymi własnościami determinującymi osobę są pewne cechy i relacje psychiczne (pamięć, charakter) lub fizyczne (ciągłość cielesna, identyczność mózgu).

Metoda myślenia krytycznego byłaby metafizycznie niemożliwa, gdyby to, kim jesteśmy, było zdeterminowane naszą charakterystyką, na którą w sposób konieczny i istotny składają się takie a nie inne stany motywacyjne. Tracąc ową charakterystykę, nie bylibyśmy tymi samymi osobami. Richard Hare odrzucał jednak powyższą koncepcję. Bronił on **nondeskryptywistycznej odmiany teorii identyczności osobowej**. Według niego, trzeba rozróżnić dwie hipotezy, z których tylko pierwsza jest logicznie niemożliwa:

⁹⁹ Taką perspektywę można także nazwać z T. Nagelem „*God eye view*” – „Boską perspektywą” czy „widokiem znikąd”. Zob. T. Nagel, *Widok znikąd*, Fundacja Aletheia, Warszawa 1997.

¹⁰⁰ Taką tezę akceptuje np. Thomas Nagel. Zob. tenże, *Subiektywność i obiektywność*, w: T. Szubka (red.), *Metafizyka...*, dz. cyt.

¹⁰¹ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 140.

¹⁰² Nagel stawia to pytanie w odniesieniu do problemu tożsamości osobowej. T. Nagel, *The Foundations...*, dz. cyt., s. 111.

¹⁰³ Zob. C. Taylor, *Review of „Freedom and Reason” by R. M. Hare*, „Mind” (74) 1965.

(1) „Jeśli Krzysztof Saja byłoby Janem Kowalskim [...]” i

(2) „Jeśli ja byłbym Janem Kowalskim [...]”¹⁰⁴.

Chociaż Krzysztof Saja i Jan Kowalski posiadają istotę – zbiór konstytuujących ich cech, to *Ja* jej nie mam (jako *Ja* właśnie, choć zgodnie z pewnymi poglądami, mogę mieć istotę jako Krzysztof Saja). Dlatego kiedy obaj używamy wyrazu „ja” z myślą o sobie, używamy go synonimicznie, jakkolwiek w różnych odniesieniach, *Ja* bowiem nie jest powiązane z żadną „istotą”¹⁰⁵.

Akceptowana przez Hare’a koncepcja tożsamości osobowej jest podobna do teorii „transcendentalnego *Ja*”, bronionej przez Zeno Vendlera¹⁰⁶. Dla Vendlera *Ja* nie jest „czymś”, o *Ja* można mówić tylko metaforycznie. Słowo „Ja” w moim akcie mowy odnosi się do Krzysztofa Saji, ale to odniesienie nie wyznacza jego znaczenia. Według Hare’a, *Ja* jest pustą ramą świadomości, obojętną na treść.¹⁰⁷ Co więcej, istnieje bardzo ważna pozytywna cecha *Ja*, która gra niebagatelną rolę w dziedzinie namysłu moralnego. Otóż **nondeskrytywistyczna teoria identyczności osobowej** głosi, że „ja” nie jest słowem do końca opisowym, lecz po części nakazowym: utożsamiając się z jakąś osobą, faktycznie albo hipotetycznie, identyfikujemy się z jej nakazami¹⁰⁸; nazywając jakąś osobę „ja”, wyrażamy zdecydowanie większą troskę o jej preferencje niż o preferencje tych osób, których tak nie nazywamy¹⁰⁹.

§ 6.5 Utylitaryzm myślenia krytycznego

Biorąc pod uwagę to, co zostało powiedziane o idealnym myśleniu moralnym należy postawić pytanie, jak należałoby zaklasyfikować stanowisko, którego przedstawicielem jest Hare’owski archanioł. Powszechnie zalicza się je do utylitaryzmu. **Utylitaryzm**, strukturalnie rzecz ujmując, jest normatywnym poglądem etycznym, będącym po-

¹⁰⁴ Por. R. M. Hare, *Comments on Vendler*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 281.

¹⁰⁵ Por. tenże, *Myślenie...*, dz. cyt., s. 153-155.

¹⁰⁶ Z. Vendler, *The Matter of Minds*, Oxford University Press 1984. Również: tenże, *Changing Places?*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 171-183.

¹⁰⁷ Zob. R. M. Hare, *Comments on Vendler...*, dz. cyt., s. 286.

¹⁰⁸ Por. tenże, *Myślenie...*, dz. cyt., s. 126.

¹⁰⁹ Zob. tamże, s. 128.

łączeniem konsekwencjalizmu, uniwersalizmu, welfaryzmu, oraz procedury sumowania i kryterium maksymalizacji¹¹⁰. Najprościej mówiąc, **konsekwencjalizm** głosi, że ocenę danego czynu, zasady, motywacji lub osoby w ostatecznym rozrachunku warunkują jedynie pozytywne lub negatywne konsekwencje, skutki, będące jej wynikiem. Wszystko, czego potrzebujemy, aby ocenić jakiś czyn, to wiedza o jego skutkach. Według **welfaryzmu**, jedynie stan dobrobytu jednostek ma wartość samą przez się – cokolwiek jest wewnętrznie dobre, cokolwiek ma nieinstrumentalną wartość, musi być wartościowe dla konkretnych osób. Drugim ważnym aspektem tego poglądu jest teza, że dobro społeczne składa się wyłącznie z dóbr indywidualnych; dobro każdej jednostki wnosi pewien wkład do dobra społecznego i nic innego się do niego nie przyczynia¹¹¹. Z tego względu ocena dobra lub zła alternatywnych stanów rzeczy, będących ewentualnymi skutkami rozważanych czynów czy reguł, dla utylitarysty musi się opierać wyłącznie na odpowiednim zbiorze wartości dla indywidualnych osób. **Uniwersalizm** utylitaryzmu zakłada, iż istotne jest dobro wszystkich osób (bądź istot czujących), **egalitaryzm** zaś, że dobro i zło każdej jednostki jest tak samo ważne. Zdaniem utylitarystów wybór słusznego działania należy dokonać ze względu na **kryterium maksymalizacji zsumowanych** dóbr jednostek: czyn jest słuszny, jeśli prowadzi do największej sumy dóbr wszystkich ludzi.

Poprzestanie jednak na ogólnym stwierdzeniu, że archanioł jest utylitarystą, byłoby poważnym uproszczeniem. Od mniej więcej półwiecza utylitaryzm jest bowiem bardzo intensywnie rozwijany i doczekał się wielu modyfikacji: niejednokrotnie odszedł daleko od swojej formy klasycznej, przypisywanej J. Benthamowi czy J. S. Millowi, tracąc niekiedy niektóre z wyżej wymienionych cech charakterystycznych. **Utylitaryzm klasyczny**¹¹² jest bezpośrednim konsekwencjalizmem czynu – stanowiskiem oceniającym jednostkowy czyn jedynie przez jego skutki – dla którego wartością istotną przy ocenie działań jest hedonistycznie interpretowane szczęście jako nadwyżka przyjemności nad przykrością

¹¹⁰ Por. S. Darwall, *Introduction* w: S. Darwall (red.), *Consequentialism*, Blackwell Publishing 2003, s. 3; R. M. Hare, *An Utilitarian Approach...*, dz. cyt., s. 156.

¹¹¹ Zob. A. Sen, *Utilitarianism and Welfarism*, w: S. Darwall (red.), *Consequentialism*, Blackwell Publishing 2003, s. 261.

¹¹² Zob. W. Sinnott-Armstrong, *Consequentialism*, w: *The Stanford Encyclopedia of Philosophy (Spring 2007 Edition)*, Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2007/entries/consequentialism/>>.

(konsekwencjalizm hedonistyczny), które to szczęście należy maksymalizować (konsekwencjalizm maksymalizujący). Ranking alternatywnych stanów świata tworzy się z bezstronnego punktu widzenia (konsekwencjalizm bezstronny) poprzez wzięcie pod uwagę wszystkich konsekwencji ocenianego działania oraz działań alternatywnych (konsekwencjalizm całościowy). Hierarchie lepszych i gorszych czynów ustala się ze względu na sumowanie przyjemności i przykrości wszystkich ludzi, traktowanych jako równoważne osoby (konsekwencjalizm uniwersalny i egalitarny).

Pierwszą różnicę między metodą myślenia krytycznego a utylityzmem klasycznym stanowi fakt, że istotną wartością, którą przy ocenie czynu musimy wziąć pod uwagę, myśląc krytycznie, nie jest przyjemność bądź przykrość, lecz spełnienie racjonalnych preferencji. Z tego względu teorię normatywną Hare'a nazwać można **utyliitaryzmem racjonalnych preferencji**¹¹³. Określić ją można również mianem **utyliitaryzmu dobrobytu** (interesu), jeśli przez interesy będziemy rozumieć motywacje wywiedzione z preferencji faktycznych przeszłych, teraźniejszych oraz przyszłych, odkrywane przez racjonalną osobę, która charakteryzuje się doskonałą wiedzą, silną wolą i niepopelnianiem błędów logicznych¹¹⁴. Mówienie o utylityzmie dobrobytu wprowadzać może jednak w błąd, gdyż Hare'a od przedstawicieli tego podejścia różni niewiara w istnienie powszechnego interesu wszystkich ludzi i uniwersalnych cech dobrobytu. Pamiętać należy, że nie chciał on zakładać żadnych materialnych wartości, które powinny być promowane; należy racjonalnie promować spełnienie roztropnych preferencji, czymkolwiek one są. Nie był więc, jak większość utylitystów dobrobytu, deskryptywistycznym naturalistą.

Hare w wielu miejscach nazywał siebie utylitystą reguł. W następujący sposób definiował dwie główne odmiany utylityzmu:

utyliitaryzm czynu [podkr. – K. S.] jest stanowiskiem głoszącym, że tzw. «zasadę użyteczności» musimy stosować bezpośrednio do indywidualnych aktów; to, co musimy zrobić, to osądzić czyn, mierząc płynący z niego oraz jego alternatyw ogólny stopień zadowolenia. **Utyliitaryzm reguł** [podkr. – K. S.], z drugiej strony, jest poglądem, że powyższy test nie powinien być stosowany do indywidualnych czynów, ale do *rodzaju* czynów. Ocena moralna czynu staje się więc procesem dwustopniowym. Czyny muszą być oceniane za po-

mocą pytania, czy są one zakazane lub pochwalane przez dane moralne prawa lub zasady¹¹⁵.

Jeśli chcemy być precyzyjni, utylityzm myślenia krytycznego faktycznie należy nazwać utylityzmem reguł, ponieważ, choć archaniot nie musi stosować się zawsze do obowiązujących zasad, każda ocena czynu polega na wyprowadzaniu uniwersalnej reguły i ocenianiu go poprzez nią. Nie możemy oceniać działania, biorąc pod uwagę jedynie aktualny stan świata. Zawsze zasada, na którą się powołamy, musi być wyrażona przy użyciu kwantyfikatorów ogólnych i własności uniwersalnych – odnosić się do więcej niż jednego przypadku. Istotną cechą klasycznie rozumianego utylityzmu czynu jest bowiem to, iż jest on konsekwencjalizmem bezpośrednim, metoda myślenia krytycznego polega natomiast na skonstruowaniu uniwersalnych reguł postępowania, poprzez które, jako archaniotowie, musimy ocenić jednostkowy czyn. Stanowisko to nazwać można **utyliitaryzmem reguł akceptowanych**, w przeciwieństwie np. do utylityzmu reguł powszechnie przestrzeganych, kulturowanych, akceptowanych publicznie czy wbudowanych w instytucje. Osobą, która powinna zaakceptować ocenianą zasadę moralną, jest idealny podmiot moralny, idealny nakazodawca – archaniot. Poznanie słuszności danego czynu stać się może naszym udziałem w pełni jedynie w idealnych warunkach epistemicznych, do których, jeśli chcemy dysponować moralną wiedzą, musimy się zbliżać.

Gdy weźmie się pod uwagę pozycję archaniota, rodzi się pytanie, czy jego namysł w praktyce nie pokrywa się jednak z bezpośrednią formą utylityzmu czynu¹¹⁶. Jeśli zdarzy się, że oceniane działanie jest związane z sytuacją, której ponowne wystąpienie jest bardzo mało prawdopodobne (cechuje ją wiele specyficznych, rzadkich własności naturalnych *RNat*, np. atak terrorystyczny z użyciem broni atomowej), ocena naszego działania w głównej mierze powinna być oparta na skutkach tego konkretnego działania związanego z praktycznie niepowtarzalną sytuacją. Co więcej, każdy przypadek, nawet jeśli dla człowieka jest podobny do innych, z perspektywy archaniota najprawdopodobniej powoduje różne dalekosiężne istotne skutki. Z tego względu Hare godził się, aby procedurę myślenia moralnego stosowaną

¹¹³ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 119.

¹¹⁴ Zob. R. E. Goodin, *Użyteczność i dobro*, w: P. Singer (red.), *Przewodnik...*, s. 286.

¹¹⁵ R. M. Hare, *Freedom...*, dz. cyt., s. 130.

¹¹⁶ Zob. tenże, *Ethical Theory and Utilitarianism*, dz. cyt., s. 224.

przez archanioła nazwać **wyrafinowanym utylitaryzmem czynu**¹¹⁷. Jeśli oprócz dwóch powyższych wersji utylitaryzmu przyjmiemy za E. Klimowicz również trzecią, tj. umiarkowany utylitaryzm czynu, można mieć uzasadnione przypuszczenie, że archanioł jest właśnie tej wersji przedstawicielem. **Umiarkowany utylitaryzm czynu**

dopuszcza możliwość kierowania się w wyborze jakąś moralną regułą, ponieważ przyjmuje, że przestrzeganie reguł zazwyczaj sprzyja szczęściu ogólnemu (albo dobru – jeśli jest to utylitaryzm idealny). Ale osoba dokonująca wyboru powinna sama zdecydować, czy w sytuacji, w której się właśnie znajduje, postępowanie zgodne z regułą będzie istotnie lepsze niż jakiegokolwiek inne możliwe do wyboru. W tym celu musi ona zastanowić się także nad konsekwencjami działania zgodnego z regułą i porównawszy wartość wszystkich możliwych czynów, zdecydować, który jest słuszny. Może się przecież zdarzyć, że czyn zgodny z regułą w danej konkretnej sytuacji nie będzie – w porównaniu z innymi – słuszny, przeto reguła przestaje nas obowiązywać¹¹⁸.

Między teorią Hare'a a klasycznym utylitaryzmem oprócz powyższych różnic istnieją także podobieństwa. Jego utylitaryzm jest uniwersalny, egalitarny i bezstronny. Zauważalne są jednak niezgodności, które w kolejnym rozdziale staną się jeszcze wyraźniejsze. Wziąwszy pod uwagę rzeczywistą epistemiczną sytuację każdego z nas, Hare zmodyfikował bowiem klasyczny utylitaryzm w jeszcze większym stopniu, niż opisaliśmy to powyżej. Idealna metoda namysłu moralnego nie może być stosowana zawsze, gdy człowiek staje przed moralnym problemem: ludzie w codziennej praktyce powinni ograniczać swoje ambicje myślenia krytycznego. Przedstawienie tych zmian stanowi treść kolejnego rozdziału.

¹¹⁷ Zob. tenże, *Utilitarianism and the Vicarious...*, dz. cyt., s. 239.

¹¹⁸ E. Klimowicz, *Utylitaryzm w etyce. Współczesne kontrowersje wokół etyki Johna Stuarta Milla*, PWN, Warszawa 1974, s. 204.

§ 7. Człowiek a myślenie krytyczne

Analizując język moralności Richard Hare doszedł do specyficznej wersji utylitaryzmu. Według niego, metoda myślenia krytycznego jest najbardziej spójna z cechami języka moralnego oraz pozwala efektywnie zmierzyć się z praktycznymi problemami etycznymi. Głównym jej bohaterem jest archanioł, idealny nakazodawca, którego sposób rozwiązywania problemów moralnych jest dla nas wzorem. Rodzi się jednak pytanie, czy teoria Hare'a nie jest dla ludzi bezużyteczna. Wymaga ona przecież od nas sprawności, których nie posiadamy. Nie jesteśmy w stanie tak jak archanioł przewidzieć wszystkich konsekwencji swoich działań ani znać chociażby konsekwencji istotnych; nie mamy wiedzy o wszystkich możliwych osobach, które wziąć musimy pod uwagę podczas namysłu moralnego; nie potrafimy jednoznacznie rozsądzić, które własności sytuacji są moralnie istotne; nie znamy wszystkich hipotetycznych sytuacji; nie potrafimy się wczuć w preferencje i interesy innych ludzi tak, aby były one naszymi itd. Wydaje się, iż człowiek nie jest zdolny za pomocą metody Hare'a dojść do żadnej efektywnej konkluzji. Jaką więc rolę przypisał Hare myśleniu krytycznemu w życiu przeciętnego człowieka? Aby odpowiedzieć na powyższe pytania, zaprezentujemy jego uzasadnienie wprowadzenia dwóch płaszczyzn myślenia moralnego, bliżej opiszemy poziom intuicyjny, omówimy zespół niezbędnych człowiekowi cnót do sprawnego myślenia krytycznego oraz rozważymy granice odpowiedzialności za zaniechanie utylitarystycznego promowania zaspokojenia preferencji.

Aby uzupełnić obraz zaprezentowany w poprzednim rozdziale, oprócz archanioła należy przedstawić jego przeciwieństwo. **Robol**¹ jest kolejną postacią zaprezentowaną w *Myśleniu moralnym*. Uosabia on sposób funkcjonowania jedynie na płaszczyźnie intuicyjnej, posługuje się wyłącznie metodą myślenia intuicyjnego. Polega ona nie na metodycznym analizowaniu danego problemu, lecz kierowaniu się oczywistymi w naszym mniemaniu zasadami, które wydają się proste, prawdziwe i niepodważalne. Robol jest niezdolny do myślenia krytycznego, tworzenia zasad, często poddaje się słabości woli. Intuicje moralne, które posiada, to dyspozycje wynikające z wychowania. Reguły domniemane przejmując od innych ludzi, a następnie traktuje je jako niepodważalne. Nie potrafi uzasadnić swoich ocen ani działać właściwie w sytuacjach wcześniej nieocenionych².

Połączeniem robola i archanioła jest **człowiek** – postać, która w niniejszym rozdziale jest przedmiotem naszego zainteresowania. Każdy z nas, w zależności od swoich uzdolnień i ilości czasu, jakim dysponuje w trakcie podejmowania decyzji, do jej rozwiązania stara się podejść raz jak archanioł (nie bacząc na zastane obyczaje, patrząc na skutki działań, ważąc preferencje i próbując maksymalizować szczęście lub minimalizować nieszczęście), innym razem zaś jak robol³ (kierując się intuicjami, tradycją czy sumieniem). To, jaką metodę i w jakich okolicznościach częściej wybieramy, zależy od naszego temperamentu, zdolności poznawczych, cech charakteru, zastanej sytuacji czy nawet światopoglądu.

Choć na pierwszy rzut oka myślenie krytyczne i intuicyjne są metodami konkurencyjnymi, na co wskazuje spór pomiędzy klasami utilitaryzmu a intuicjonizmu, dla Hare'a stanowią one część jednej struktury – dwupoziomowego utilitaryzmu reguł – w której owe płaszczyzny pełnią różne funkcje. Człowiek powinien starać się używać obu metod w zależności od sytuacji oraz swoich zdolności. Gdy mamy do rozwiązania jakiś konkretny problem moralny, gdy

zastanawiamy się, jakie działanie byłoby moralnie słuszne, na pierwszym etapie powinniśmy się odwołać do myślenia intuicyjnego, które podpowiada nam pewną zasadę domniemaną. Jeśli nie budzi ona naszych wątpliwości lub nie mamy czasu na namysł krytyczny, powinniśmy postąpić zgodnie z zaleceniem, które ona wyraża. Większość naszych codziennych problemów moralnych może zostać adekwatnie rozwiązana już na tym etapie. Zdarzają się jednak sytuacje wyjątkowe lub nowe. W takim wypadku, gdy nie jesteśmy przekonani co do zasady domniemanej lub intuicja podpowiada nam różne sprzeczne reguły oraz mamy czas i czujemy się intelektualnie i moralnie na siłach aby myśleć krytycznie, powinniśmy poszukać zasady krytycznej, którą zaleciłby archanioł. Następnie, jeśli uda się nam przeprowadzić namysł krytyczny, którego rezultatem jest zasada krytyczna, powinniśmy się zastanowić, czy nie byłoby warto ją uprościć i, jeśli istnieje taka możliwość, przekazać innym. Warto konstruować bowiem nowe zasady domniemane dla nieznanymi wcześniej sytuacji, jeśli prawdopodobieństwo ich ponownego wystąpienia jest znaczne. Będą one użyteczne dla nas samych, naszych dzieci i wszystkich, którym jesteśmy je w stanie zalecić. W innych, bardziej wyjątkowych przypadkach podejmowanie wysiłku adekwatnego upraszczania nie jest wystarczająco uzasadnione.

Powyższą ścieżkę rozwiązywania problemu moralnego schematycznie przedstawić można w następujący sposób:

¹ Słowo „robol” jest tłumaczeniem angielskiego słowa *prole*. Hare używa tej nazwy, odwołując się do osób zaliczanych do najniższej klasy społeczeństwa Oceanii, którą opisał George Orwell w powieści *Rok 1984*.

² Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 61-85. Archanioł i robol są karykaturalnymi postaciami z najwyższej i najniższej klasy społecznej Platońskiego państwa. Zob. tamże, s. 62.

³ Zob. tamże, s. 62.



Rysunek 11. Droga rozwiązywania problemów moralnych wg R. M. Hare'a

Zdaniem Hare'a, choć myślenie krytyczne jest epistemicznie i logicznie nadrzędne, człowiek powinien stosować je z dużą ostrożnością. W większości przypadków ludzie powinni się kierować zasadami domniemanymi. Duży nacisk na stosowanie zasad domniemanych, widoczny w tekstach Hare'a, uzasadniony jest próbą jak najpełniejszego zbliżenia skutków ocen wydawanych przez niedoskonałego człowieka do rezultatów stosowania myślenia krytycznego przez idealnego nakazodawcę. Gdyby archanioł, który myśli krytycznie, wiedział, że musi stworzyć reguły myślenia moralnego dla ludzi, z racji ludzkich wad zaleciłby im podporządkowanie się prostym zasadom domniemanym. Hare wprowadził do swojego systemu płaszczyznę intuicyjną,

ponieważ uważał, iż dla człowieka prosty utilitaryzm czynu jako powszednia metoda myślenia moralnego jest nieużyteczny⁴. Człowiek ma zbyt wiele epistemicznych i moralnych wad, aby stosowanie przez niego takiej metody przyniosło pożądane efekty w postaci maksymalizacji czy promowania spełnienia racjonalnych preferencji. Poza przypadkami bardzo wyjątkowymi, nie powinien kusić się on na myślenie krytyczne⁵. Jedynie taka jego postawa efektywnie przyczynia się do zbliżania się do rezultatów, które osiągnąłby archanioł, przestrzegający zasad krytycznych. Jeśli będziemy stosowali myślenie intuicyjne, nasze działanie prawie zawsze będzie **moralnie racjonalne**, tzn. **maksymalizujące oczekiwaną użyteczność**⁶. Celem pedagogiki moralnej jest zatem wychowanie ludzi do jak najbardziej efektywnego i skutecznego przestrzegania przez nich zasad, które wybrałby dla nich archanioł. Jak pisał Hare, myślenie krytyczne jest przydatne wtedy, gdy filozofujemy, myślenie intuicyjne zaś gdy jesteśmy w moralnych kłopotach⁷.

Duży nacisk położony na posługiwanie się myśleniem intuicyjnym wynika również z faktu, że Hare doceniał wagę tradycji i przekazywanych w jej imieniu prawideł. Są one niezbędne do funkcjonowania człowieka w świecie. Najcenniejszym skutkiem powstania języka moralnego są bowiem reguły postępowania. Stwarzane na bazie naszego doświadczenia i przekazywane z pokolenia na pokolenie, stanowią praktyczne instrukcje, ułatwiające ludziom poruszanie się w społeczeństwie⁸. Aby wiedza moralna mogła być kumulowana, założyć jednak trzeba, że określone zasady moralne stosują się do wielu sytuacji oraz mogą być proste na tyle, iż łatwo je zapamiętać i ich nauczać. Jeśli bowiem chcemy przekazywać wiedzę o regułach postępowania, nie możemy do każdej sytuacji podchodzić jak do indywidualium ani tworzyć zbyt skomplikowanych zasad moralnych. Proces edukacji moralnej odbywa się bowiem już w wieku dziecięcym.

Nie istnieje jednak jednoznaczna odpowiedź na pytanie, jakim stopniem szczegółowości powinny się cechować zasady domniemane. Bardzo dużo zależy od konkretnych ludzi i ich grup: od tego, na ile potrafią je pojąć i ograniczać się w myśleniu krytycznym. Zależ-

⁴ Zob. tamże, s. 51-53, 154.

⁵ Zob. tenże, *Principles*, w: tenże, *Essays in Ethical...*, dz. cyt., s. 65.

⁶ Zob. tamże, s. 63-64.

⁷ Zob. tamże, s. 66.

⁸ Zob. tenże, *The Language...*, dz. cyt., s. 60-68.

nie od możliwości przyjęcia i używania bardziej złożonych zasad domniemanych możemy rozróżnić podgrupy w społeczeństwie, i poszczególne jednostki w tych podgrupach, a także różne stadia rozwoju poszczególnych jednostek. Dla większości ludzi bowiem stosowane przez nich zasady domniemane będą się stawały coraz bardziej złożone w miarę wzrostu ich życiowego doświadczenia – przebywania w wielu różnych sytuacjach wymagających moralnych wyborów⁹. Wraz z dorastaniem, dzieci uczą się coraz większej ilości wyjątków – zasady moralne stają się dla nich bardziej szczegółowe; uczą się kolejnych reguł, które odsłaniają im odmienny aspekt ocenianej sytuacji. Na pewnym etapie rozwoju ludzie powinni poznać jednak metodę myślenia krytycznego, aby w skomplikowanych, nowych życiowych sytuacjach nie pozostawać bezbronnymi wobec sprzecznych intuicji moralnych oraz aby umieć racjonalnie uzasadnić przyjmowane reguły domniemane. Z perspektywy trzeciej osoby nie jesteśmy jednak w stanie podać kryteriów, kiedy ktoś powinien myśleć krytycznie, a kiedy stosować zasady domniemane¹⁰. Tylko indywidualna ocena własnych możliwości może dostarczyć nam tej odpowiedzi. Ludzie z doświadczeniem życiowym, którzy biegle rozumują, umieją odpowiednio stosować obie z nich¹¹. „Musimy poznać siebie, żeby wiedzieć, czy możemy zaufać sobie na tyle, by grając archaniola, nie skończyć w miltonowskim obozie jako upadły anioł”¹².

§ 7.1 Czy utilitaryzm jest zbyt wymagający?

Jeśli znajdziemy się w sytuacji, w której powinniśmy podjąć myślenie krytyczne, czy jako ludzie będziemy w stanie je udźwignąć? Jeden z najważniejszych zarzutów skierowanych w stronę utilitaryzmu głosi, iż utilitaryzm jest zbyt wymagający. Tezę tę można interpretować na dwa sposoby: epistemicznie i moralnie. Według pierwszej interpretacji, utilitaryzm wymaga od nas posiadania niemożliwej do osiągnięcia wiedzy i umiejętności poznawczych: mamy znać preferencje innych ludzi, traktować je jako nasze własne, przewidywać skutki działań,

⁹ Zob. tenże, *Ethical Theory and Utilitarianism*, dz. cyt., s. 225.

¹⁰ Zarzut, który poddaje w wątpliwość istnienie kryterium dotyczącego tego, kiedy myśleć intuicyjnie, a kiedy krytycznie, można znaleźć w: D. A. J. Richard, *Prescriptivism...*, dz. cyt., s. 123.

¹¹ Zob. R. M. Hare, *Myślenie...*, dz. cyt., s. 63.

¹² Tamże.

potrafić zsumować spełnienie preferencji i ustawić te sumy w hierarchii, znać wartość alternatywnych dróg wyboru. Wysuwając taki zarzut, zasadę użyteczności traktuje się jako kryterium decyzyjne, które w praktyce jest nieużyteczne. Przy drugiej interpretacji twierdzi się, że utilitaryzm jest zbyt wymagający, gdyż normalny człowiek, który miałby spełnić utilitarystyczne żądania, codziennie powinien czuć się winny. Bardzo rzadko bowiem maksymalizujemy szczęście ogółu: nasze zasoby przeznaczamy na dostarczenie przyjemności sobie i swoim najbliższym, a nie tym, którzy najbardziej tego potrzebują. Gdy przyjmujemy maksymalizację szczęścia jako kryterium słuszności czynu, okaże się – głosi ta interpretacja – że przytłaczająca większość naszych czynów w myśl utilitarystycznych standardów jest moralnie niesłuszna¹³.

§ 7.1.1 Czyn zły a czyn niesłuszny

Odpowiedzią Hare’a na powyższe zarzuty było rozdzielenie trzech elementów, których tożsamość zakłada się przy wysuwaniu podanych zastrzeżeń: definicji czynu moralnie słusznego, kryterium podejmowania moralnych decyzji oraz kryterium pociągania kogoś do moralnej odpowiedzialności¹⁴. Związane są one z podziałem na myślenie krytyczne i intuicyjne. Myślenie krytyczne, gdy weźmie się pod uwagę cechy człowieka i otaczającego go świata, ma bowiem tendencję do samoograniczania się, co wyjaśniliśmy, przedstawiając uzasadnienie dla myślenia intuicyjnego. Istnieją bardzo dobre utilitarystyczne powody, dla których ludzie powinni być zachęceni do ograniczonego dbania o maksymalizowanie użyteczności. Myślenie krytyczne nie jest dobrym kryterium podejmowania decyzji, biorąc bowiem pod uwagę nasze wady, większą korzyść przyniesie przestrzeganie prostych zasad domniemanych niż myślenie krytyczne podejmowane na własną rękę. Z tych samych powodów dla utilitarysty jest ważne, abyśmy dbali o rozwój w ludziach dyspozycji, uczuć, cech charakteru, motywacji, przekonań i postaw, które umożliwiają im przestrzeganie i kul-

¹³ Tim Mulgan uważa ten zarzut za podstawę wielu innych ataków skierowanych w stronę utilitaryzmu i, traktując go jako bardzo istotny, poświęcił mu całą książkę. Zob. T. Mulgan, *The Demands...*, dz. cyt.

¹⁴ Zob. R. M. Hare, *Utilitarianism and the Vicarius...*, dz. cyt., s. 239-241; tenże, *Comments on Scanlon*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 261.

tywowanie tych zasad¹⁵. Jedynie bowiem intensywne uczucia moralne i trwałe nawyki są w stanie skutecznie uchronić człowieka od zbyt pochopnego myślenia krytycznego.

Kryterium podejmowania decyzji nie musi wynikać z definicji czynu słusznego. Zdaniem Hare'a, czyn jest **moralnie słuszny**, jeśli pokrywa się z tym, co zrobiłby idealny nakazodawca, który myśli krytycznie. Nie jest jednak prawdą, że człowiek ma obowiązek zawsze dokonywać rzeczy słusznych. Mamy obowiązek wybierać czyny dobre. **Dobry czyn** jest tym, co zrobiłby dobry człowiek. Natomiast **dobry człowiek**, człowiek o moralnie dobrym charakterze, jest osobą, która posiada bardzo mocno zinternalizowane zasady domniemane. Zasady takie w większości sytuacji zaprowadzą go do działań słusznych, optymalizują one bowiem prawdopodobieństwo ich uczynienia¹⁶. Człowiek moralnie dobry nie waha się np., czy powinien uwięzić żonę swego sąsiada, czy nie. Osoba taka nie potrafi złamać domniemanych zasad moralnych bez poczucia winy, czuje się zawsze źle, gdy sprzeniewierza się moralnym zasadom domniemany. Celem moralnej edukacji jest bowiem między innymi wyrobienie w dzieciach trwałych motywacji: uczuć, dyspozycji do działania, postaw, które sprzyjają przestrzeganiu zasad domniemanych oraz nie mogą być lekceważone bez poważnych wyrzutów sumienia¹⁷. Występowanie moralnych uczuć, takich jak wyrzuty sumienia czy moralne oburzenie, jest zazwyczaj symptomem dobrego wychowania.

To, że działanie dobrego człowieka nie zawsze okaże się słuszne, wynika przede wszystkim z dwóch powodów: po pierwsze, nie zawsze jest on w pełni poinformowany, posiada poznawcze ograniczenia i zdając sobie z tego sprawę, woli trzymać się zasad domniemanych; lub, po drugie, nawet jeśli ma uzasadnione wątpliwości co do słuszności danego czynu, który jest zgodny z zasadą domniemaną, nie potrafi postąpić inaczej niż w sposób, w jaki został wychowany lub do jakiego skłaniają go silne moralne emocje ściśle powiązane z akceptowanymi przez niego zasadami domniemany. Zdaniem Hare'a, sytuacje, w których dobry człowiek postąpi niesłusznie, są jednak rzadkie¹⁸.

¹⁵ Zob. tenże, *Utilitarianism and the Vicarius...*, dz. cyt., s. 239-241; W. H. Shaw, *Contemporary Ethics. Taking Account of Utilitarianism*, Blackwell Publishers, Oxford 1999, s. 170.

¹⁶ Zob. R. M. Hare, *Ethical Theory and Utilitarianism*, dz. cyt., s. 228.

¹⁷ Zob. tenże, *Principles*, dz. cyt., s. 63-64.

¹⁸ Zob. tenże, *Ethical Theory and Utilitarianism*, dz. cyt., s. 227.

Również **moralna wina** nie powinna pokrywać się z działaniem słusznym lub niesłusznym. Nie należy moralnie karcić osoby, która postąpiła niesłusznie, jeśli działała w myśl zasady domniemanej. Fakt niepociągania do odpowiedzialności osób, które nie maksymalizują użyteczności, można wyjaśnić na podstawie prawd dotyczących psychiki ludzi oraz utylityzmu. Jeśli moralność byłaby bowiem zbyt surowa, ludzie zniechęcili by się do jej przestrzegania, co ograniczałoby możliwą do uzyskania użyteczność.

§ 7.1.2 Obowiązki supererogacyjne i integralność ludzkiego życia

Kolejny, często powtarzany argument przeciwko utylityzmowi, który we wpływowy sposób podniósł J. Rawls, głosi, że „utyliaryzm ignoruje niezależność osób”¹⁹, utylityści bowiem przedkładają interes zbiorowości nad preferencje jednostki. Jego egzemplifikacją są hipotetyczne sytuacje, w których, według krytyków, utylitysta zalecałby poświęcenie jednej czy kilku osób dla maksymalizacji szczęścia ogółu. Możemy wyobrazić sobie utylitystyczne żądanie, aby w sytuacjach gdy brakuje ratujących życie organów do przeszczepu, pobierać je od osób, które mają je po parze. Co więcej, moglibyśmy głosić, że ponieważ wykorzystanie wszystkich organów jednego człowieka może uratować wielu innych, mamy obowiązek poświęcić swoje życie i rozdać swoje ciało.

Zarzut, że utylityzm ignoruje niezależność osób, nie byłby jednak przekonujący, jeśli mielibyśmy czytać go dosłownie²⁰. Moralność ujęta w jakimkolwiek systemie etycznym wymaga bowiem od ludzi poświęcenia własnych interesów dla korzyści innych. Nie wolno nam kłamać, zadawać bólu, kraść; mamy obowiązek niesienia pomocy osobom, które uległy wypadkowi drogowemu, również wtedy gdy nie jest nam to na rękę. Aby nadać mu bardziej adekwatny charakter, należałoby go przeformułować: utylityzm ignoruje niezależność osób, gdyż wymaga on zbyt daleko idącego poświęcenia jednostki dla dobra innych lub ogółu. W takiej formie argument Rawlsa redukuje się do moralnej wersji zastrzeżenia, iż utylityzm jest zbyt wymagający²¹. Zgodnie z tą krytyką, utylityzm czyni powinnością działania, które choć w powszechnej

¹⁹ Zob. J. Rawls, *Teoria sprawiedliwości*, dz. cyt., s. 43.

²⁰ Zob. W. H. Shaw, *Contemporary Ethics...*, dz. cyt., s. 128.

²¹ Zob. T. Mulgan, *The Demands...*, dz. cyt., s. 17-18.

opinii mogą być działaniami chwalebnyymi, nie są i nie mogą być jednak obowiązkowe. Zasada maksymalizacji szczęścia powszechnego kosztem naszego życia może być regułą supererogacyjną, chwalebną, nadobowiązkową, lecz nie naszą powinnością.

Inną cechą utylityzmu, która powoduje, że teoria ta jest zbyt wymagająca, jest przyjmowane przez jego zwolenników założenie, iż różnica między działaniem a zaniechaniem działania nie jest moralnie istotna. Zdaniem Hare'a, na poziomie myślenia krytycznego, nie ma istotnej różnicy między zabiciem a zaniechaniem pomocy ze skutkiem śmiertelnym; zaniechanie jest też działaniem, za które jesteśmy odpowiedzialni ze względu na jego konsekwencje²². Wyniki przyjęcia takiego założenia w niektórych przypadkach mogą być jednak kontrintuicyjne: jesteśmy w taki sam sposób winni śmierci osób, które bez naszej pomocy umierają w krajach trzeciego świata z pragnienia, głodu i chorób, jak osób, które uśmiercilibyśmy, wbijając im nóż w plecy. Utylityzm stawia więc wymagania, które są nierozsądne.

Zdaniem Hare'a, obrońca jego wersji utylityzmu nie ma jednak problemów z uzasadnieniem zaleceń, które nie są obowiązkami, a które mają charakter nadobowiązkowy. Zauważyć bowiem należy, że argument powyższy opiera się na przekonaniach z płaszczyzny intuicyjnej: głosi, iż intuicyjne zasady domniemane nie wymagają od nas tego, czego żądałby namysł krytyczny. Jeśli jednak rozdzielimy dwa poziomy myślenia moralnego i zaakceptujemy fundamentalną wagę zasad domniemanych dla praktyki moralnej zwykłego człowieka, problem zaleceń nadobowiązkowych może zostać wyjaśniony. Co więcej, Hare wyróżnił nie tylko zasady krytyczne i domniemane, lecz również pewne podpoziomy tych drugich. Błędem byłoby bowiem sądzić, że między zasadami domniemanymi, które muszą być wspólne niemal wszystkim, a szczegółowymi zasadami krytycznymi nie występuje żaden poziom pośredni²³. Istnieją

przynajmniej trzy podpoziomy zasady domniemanych: te, które są wspólne dla wszystkich, te, które są wspólne dla poszczególnych ról, oraz zasady osobiste właściwe jednostkom. Nietrudno w tych kategoriach przedstawić ujęcie mechanizmów supererogacji. Pospolite zasady domniemane, wspólne wszystkim, oraz zasady właściwe poszczególnym rolem pobudzają to, co w konwencjonalnym ujęciu nosi nazwę obowiązków czy też w szerszym znaczeniu, zobowiązań²⁴.

²² Zob. R. M. Hare, *Abortion...*, dz. cyt., s. 161.

²³ Tenże, *Myślenie...*, dz. cyt., s. 248.

²⁴ Tamże, s. 251.

Zasady domniemane rozbić można więc na klasę reguł, które muszą być przestrzegane powszechnie, klasę reguł związanych z pełnionymi rolami społecznymi (np. kodeksy zawodowe) oraz indywidualne zasady domniemane, ufundowane na jednostkowych cechach danych osób.

Jeśli zastanawiamy się nad zakresem zasad nadobowiązkowych, które powinny stać się naszym indywidualnym drogowskazem, każdy z nas powinien zadać sobie pytanie, jaka jest miara świętości, do której jest zdolny²⁵. Zasady domniemane, zarówno powszechne, jak i zawodowe czy indywidualne, nie mogą być jednak zbyt wymagające ze względu na psychiczne i fizyczne ułomności człowieka: mamy wrodzoną skłonność do dbania bardziej o interes własny i naszych najbliższych, niż osób obcych.

Moralność o najwyższej akceptowalności-użyteczności musi nadawać się do zastosowania w praktyce. Tymczasem większości z nas z trudem przychodzi żyć konsekwentnie w zgodzie nawet z tymi pospolitymi zasadami, zwłaszcza gdy jesteśmy zmęczeni, przestraszeni albo bardzo źli²⁶.

Zbyt wiele szkód przyniosłoby ganienie ludzi za to, że nie przestrzegają oni nadmiernie wymagających reguł domniemanych. Z tego względu skuteczniejsze jest zalecanie dobrowolnych zasad supererogacyjnych, chwalenie za ich przestrzeganie, nagradzanie i stawianie pomników bohaterom moralności niż próba powszechnego wymuszania stosowania zasad, które są zbyt surowe. Intuicyjną wagę rozróżnienia między działaniem a zaniechaniem możemy natomiast zrozumieć, jeśli zauważymy, że np. większość przypadków zabicia drugiego człowieka różni się pod wieloma innymi istotnymi względami od przypadków nieudzielenia pomocy ze skutkiem śmiertelnym. Czyny pierwszego rodzaju z dużo większym prawdopodobieństwem będą złe niż czyny rodzaju drugiego i dlatego jest racjonalne, aby ludzie zaszczepili w sobie dyspozycje i reguły, które surowiej potępiają aktywne czynienie krzywdy niż nieudzielenie pomocy z podobnym skutkiem²⁷.

Z powyżej przedstawionym zarzutem powiązany jest kolejny, tzw. argument z integralności ludzkiego życia. Klasyczne jego sformułowanie zawdzięczamy Bernardowi Williamsowi²⁸, który zadał pytanie, jak utylityrystyczny podmiot, nie burząc sensowności, integralności, nie-

²⁵ Zob. tamże, s. 249.

²⁶ Tamże, s. 252.

²⁷ Tenże, *Abortion...*, dz. cyt., s. 161.

²⁸ B. Williams, *Utilitarianism For and Against*, Cambridge University Press, Cambridge 1973, s. 116.

zależności swojego życia, może traktować siebie jedynie jako jeden z wielu podmiotów i podejmując decyzje, nie ujmować swoich pragnień, preferencji, wartości, zasad jako wyróżnionych, lecz zważać na ich siłę z bezstronnego punktu widzenia. Podobny zarzut przedstawił Peter Railton²⁹, posługując się pojęciem alienacji. Jego zdaniem, utilitaryzm, żądając od nas zawsze przyjmowania bezstronnej perspektywy, alienuje nas od własnego życia. Jeśli sensowne życie zdefiniujemy jako czynne zaangażowanie w rzeczy dla nas wartościowe, jako próbę realizacji ważnych dla nas projektów, może się okazać, że życie idealnego, utilitarystycznego podmiotu nie będzie miało sensu. Utilitarysta bowiem, podejmując jakiegokolwiek decyzje, nie powinien traktować siebie jako kogoś ważniejszego od innych: pragnienia swoje i innych ludzi musi traktować jako równoważne. Argument ten głosi więc, że utilitaryzm nie ceni wystarczająco poważnie jedyności i integralności życia indywidualnego, z ich ambicjami, planami, marzeniami, wartościami i celami, każąc im dystansować się wobec samych siebie. Wymaga, abyśmy własne plany traktowali jako równoważne z celami innych i bezstronnie spełniali te preferencje, które są najsilniejsze, lub czynili to, co maksymalizuje spełnienie preferencji zbiorowości.

Na powyższy zarzut odpowiedzieć można, wskazując, po pierwsze, że również wiele alternatywnych wobec utilitaryzmu teorii etycznych, np. chrześcijański humanizm wraz z główną jego wartością, jaką jest miłość miłosierna (*agape*), w myśl standardów Williama zbytnie daleko ingeruje w integralność indywidualną³⁰. Po drugie, warto wskazać na dychotomię dwu poziomów zasad domniemanych i nadobowiązkowość niektórych reguł. Jeśli bowiem na płaszczyźnie intuicyjnej moralność nie wymaga od człowieka rezygnacji z jego własnych dążeń w celu maksymalizacji użyteczności, a obowiązkowy charakter mają jedynie minimalne reguły powszechne i te, które są związane z pełnionymi rolami społecznymi, co pokazaliśmy powyżej, omawiając zasady supererrogacyjne, to w ramach tak pojętych granic człowiek posiada wystarczająco dużo swobody, aby móc się realizować. Na płaszczyźnie intuicyjnej ilość moralnych ograniczeń, które nadają kształt naszemu życiu, nie jest dużo, i każdy z nas w ich ramach może osiągać również swoje indywidualne cele i wartości.

²⁹ P. Railton, *Alienation, Consequentialism and Morality*, „Philosophy and Public Affairs” (13) 1984, s. 134.

³⁰ Zob. R. M. Hare, *Ethical Theory and Utilitarianism*, dz. cyt., s. 219.

§ 7.1.3 Utylitaryzm a stronniczość

Z moralną interpretacją zarzutu, że utilitaryzm jest zbyt wymagający, w istotny sposób związana jest kolejna wątpliwość, z którą Hare musiał się borykać. Według Williama D. Rossa³¹, utilitaryzm nie potrafi wyjaśnić i uzasadnić, dlaczego w wielu sytuacjach powinniśmy być stronniczy. Obowiązek patriotyzmu, przedkładania interesów swojej rodziny, swoich przyjaciół nad interesy osób obcych wymyka się bezstronnej kalkulacji utilitarystów. Utylitaryzm jest zbyt wymagający, gdyż nakazuje nam spełniać silniejsze preferencje osób obcych, a nie słabsze, ale ważniejsze dla nas i moralnie nadrzędne, preferencje naszych dzieci i przyjaciół. Ktoś, kto podejmuje swoje decyzje na podstawie bezstronnej oceny preferencji różnych osób, musi być fatalnym przyjacielem, mężem, żoną, rodzicem czy członkiem zespołu sportowego. Czy nie jest bowiem tak, iż wszystkie te role wymagają od nas stronniczości?

W wielu wypadkach odkrycie uniwersalności i bezstronności ocen moralnych należy traktować jako niewątpliwy postęp w dziedzinie moralności. Nieistotność majątku, pochodzenia, koloru skóry, płci czy wyznania dla oceny czynu lub człowieka jest wszak bardzo cenioną dziedziną oświeceniową zdobyczą. Ludzie we wcześniejszych wiekach często przejawiali zbyt daleko posuniętą stronniczość: trybalizm, nepotyzm, rasizm, seksizm i nacjonalizm dla wielu osób naszej kultury wydają się dziś błędnymi ścieżkami w dziejach moralności. Z drugiej jednak strony, zdaniem Hare'a, istnieją sytuacje, w których stronniczość powinna być zalecana i są one, wbrew krytykom, w najwyższym stopniu spójne z metaetycznymi założeniami i przedstawioną metodą myślenia moralnego. Choć uniwersalność, która jest podstawą bezstronności sądów moralnych, jest logicznym prawem języka moralnego, to w wielu uzasadnionych przypadkach bezstronny namysł moralny może prowadzić do preskrypcji, która zaleca stronniczość. Balans między odpowiednią ilością materialnej stronniczości, odpowiednim przedmiotem działań stronniczych a zalecaną bezstronnością w relacjach między ludźmi jest trudny do ustalenia i powinien być rozważony dla różnych rodzajów sytuacji i relacji oddzielnie. Rozwiązań należy poszukiwać, po pierwsze, wśród zasad domniemanych lub, po drugie, gdy mamy do czynienia z trudniej-

³¹ W. D. Ross, *The Right and the Good*, Oxford University Press, Oxford 1930, s. 19.

szymi i kontrowersyjnymi przypadkami, rozumując jak archanioł. Myśląc krytycznie, należy zadać sobie pytanie, np. czy bardziej chcielibyśmy, stawiając się na miejscu zaangażowanych osób, aby obowiązywała zasada: „W pierwszym rzędzie powinno się przyjąć do pracy swoich najbliższych”, czy też: „Rekrutacja kadry pracowniczej powinna mieć charakter bezstronnego konkursu”. Archanioł najprawdopodobniej dla wielu takich problemów przedstawiłby bardzo szczegółowe, wielowątkowe zalecenie, które wykracza poza prosty, przedstawiony powyżej dualizm. Wiele zależy bowiem od sytuacji zaangażowanych osób, ich wzajemnych relacji, kontekstu społecznego, ekonomicznego, politycznego czy ideologicznego, który przekłada się na siłę i racjonalność wchodzących w grę preferencji.

Domniemane zasady głoszące przykładanie większej wagi do interesów naszej rodziny, przyjaciół i kolegów niż osób nam obcych są tymi, które w wielu wypadkach zaleciłby człowiekowi archanioł, znając jego potrzeby i pragnienia. Konieczność posiadania rodziny, silna potrzeba posiadania kręgu przyjaciół i znajomych oraz to, iż relacje te z konieczności bazować muszą na stronniczych stosunkach, są bowiem faktami, które archanioł utylityrysta musi wziąć pod uwagę, wydając uniwersalne sądy moralne. Z tego względu każdy, kto myśli krytycznie i jest racjonalny, gdyby postawił się na miejscu innych ludzi, silniej chciałby, aby obowiązywała podobna zasada do „Przedkładaj potrzeby swoich najbliższych nad potrzeby osób ci obcych” niż „W równy sposób staraj się zaspokajać potrzeby każdej osoby”. Wydaje się pewne, że archanioł, biorąc pod uwagę siłę preferencji, uniwersalnie zalecałby np. przedkładanie interesów własnych dzieci nad interesy dzieci obcych oraz, biorąc pod uwagę obecny kontekst społeczno-ekonomiczno-polityczny krajów Europy, niekierowanie się przez nauczycieli osobistymi sympatiami przy wystawianiu ocen.

§ 7.2 Utylityzm a sprawiedliwość

Z powyżej rozważanym pojęciem bezstronności i równego traktowania związane jest pojęcie sprawiedliwości. J. Rawls, w dziele *Teoria sprawiedliwości*³² postawił utylityrystom zarzut, że nie potrafią wyjaśnić oni idei sprawiedliwości społecznej: utylityryzm, chcąc maksymalizować użyteczność, nieczuły jest bowiem na rozkład dóbr

³² Zob. J. Rawls, *Teoria sprawiedliwości...*, dz. cyt., s. 41.

wśród danej grupy. Zalecałby on niesprawiedliwe, nierówne, wyzyskujące modele organizacji społeczeństwa, jeśli maksymalizowałyby użyteczność. Argument ten przybrał znaną formę zarzutu, iż utylityryzm w pewnych okolicznościach zalecałby niewolnictwo jako praktykę maksymalizującą szczęście ogółu.

Richard Hare w artykule *That is wrong with slavery*³³ bronił swojej teorii i powyższy zarzut uznał za nieuzasadniony. Zaproponował rozważenie argumentu myślowego o dwóch wyspach, które ogłosiwszy niepodległość w dziewiętnastym wieku, rozpoczęły organizację życia społecznego według dwu odmiennych modeli. Mieszkańcy Juby, pierwszej wyspy, utrzymali instytucję niewolnictwa, poddając ją jednak pod kontrolę państwa, co w konsekwencji przyczyniło się do dużo bardziej humanitarnego sposobu podchodzenia do niewolników: ograniczenia godzin ich pracy czy eliminacji okrutnego traktowania. Juba dzięki wprowadzonym reformom i niewolnictwu stała się bardzo dobrze prosperującą wyspą. Mieszkańcy drugiej krainy – Kamajki – nie zgodzili się na istnienie niewolnictwa, przez co Kamajka popadła w stagnację i nie rozwijała się tak prężnie jak Juba. Hare szczegółowo opisał warunki panujące na dwóch wyspach, aby przedstawić jak najbardziej realny przykład, potwierdzający wątpliwości swoich krytyków. Następnie zarysował dwie strategie, które można powziąć, aby odeprzeć rozważany zarzut.

Po pierwsze, utylityrysta może zaakceptować zaproponowane fakty i zgodzić się, że istnieją sytuacje, w których niewolnictwo mogłoby być zalecane przez idealnego nakazodawcę. Wolność zazwyczaj prowadzi do szczęścia czy spełnienia preferencji, mogą istnieć jednak sytuacje, w których nie będzie to prawdą. Wolność nie jest najwyższą wartością i może być poświęcona w imię szczęścia – spełnienia racjonalnych preferencji. W czasach ataków terrorystycznych ludzie często gotowi są rezygnować z wolności w zamian za wyższy poziom bezpieczeństwa. Możliwość istnienia moralnie słusznej instytucji niewolnictwa nie podważa jednak, zdaniem Hare’a, racjonalności posiadanej przez nas zasady domniemanej, iż nikt nie powinien być niewolnikiem. Sytuacja, która mogła zdarzyć się na dwóch wyspach, jest bardzo mało prawdopodobna lub zgoła nierealna, a takie przypadki, jak już wspomnieliśmy, nie są istotne dla

³³ Zob. R. M. Hare, *What is wrong with slavery*, „Philosophy and Public Affairs” (2) 1979, s. 103-121.

namysłu moralnego. Współczesna niechęć do niewolnictwa jest spowodowana historycznym doświadczeniem tej instytucji: istniejące dotychczas systemy niewolnicze zawsze przyczyniały się bowiem do nieuzasadnionej ilości cierpienia. Choć wniosek taki opiera się jedynie na indukcji i nie ma on konkluzywnego charakteru, stanowi wystarczającą rację, aby walczyć z niewolnictwem.

Po drugie, można wątpić, że bogactwo lub bieda Juby i Kamajki zostały spowodowane instytucją niewolnictwa. Dlaczego bowiem to właśnie ta cecha przyczynia się do zamożności Juby? Może swą przyczynę ma ono gdzie indziej? Czy nie możemy wyobrazić sobie państwa o podobnych zyskach, które zostałyby osiągnięte dzięki wolnej, najemnej pracy? Powodem dobrobytu znanych nam najbogatszych państw na świecie nie jest niewolnictwo. Zdaniem wielu ekonomistów, są nimi między innymi innowacyjność i przedsiębiorczość, w znacznym stopniu zależna od wolności jak najszerzych grup ludzi.

Według Hare'a, pojęcie sprawiedliwości sprowadza się do nadania równej wagi interesom wszystkich ludzi – do bezstronności będącej podstawą myślenia krytycznego³⁴. Owa idea sprawiedliwości, nosząca niekiedy nazwę **sprawiedliwości formalnej**, jest główną cechą zasady użyteczności. Tak jak nie istnieje jeden zalecany wszystkim ludziom poziom stronniczości, nie istnieje również jedna słuszna, sprawiedliwa metoda dystrybucji dóbr, jeden sposób podziału danych wartości dla wszystkich społeczeństw. Podział taki może zostać wyznaczony jedynie wtedy, gdy weźmiemy pod uwagę istotne fakty, racjonalne preferencje oraz przeprowadzimy preferencje przez trzy etapy uniwersalizacji. Słuszny, czy w szerokim sensie sprawiedliwy, podział dóbr przebiega według zasady, którą zaakceptowałby dla członków danego społeczeństwa myślący krytycznie archanioł.

Choć sprawiedliwość formalna nie prowadzi z konieczności do równości w rozdziale mienia³⁵, istnieją jednak bardzo dobre utylitarystyczne racje, aby do niej dążyć. Głównym powodem jest znana w ekonomii zasada malejącej użyteczności krańcowej (pierwsze prawo H. H. Gossena). Głosi ona, iż zadowolenie osiągnięte ze spożycia każdej następnej jednostki danego dobra zmniejsza się aż do zaspo-

³⁴ Zob. tenże, *Ethical Theory and Utilitarianism*, dz. cyt., s. 216; tenże, *Myślenie...*, s. 196-200.

³⁵ Zob. tenże, *Ethical Theory and Utilitarianism*, dz. cyt., s. 216.

kojenia na nie potrzeby. Reguła ta ma istotne znaczenie dla utylitarystycznego namysłu nad sposobem dystrybucji dóbr³⁶. Jeśli na przykład posiadanie czterech samochodów nie jest cztery razy bardziej satysfakcjonujące niż posiadanie jednego, to w konsekwencji, jeśli bierzemy pod uwagę cztery osoby, świat, w którym każda osoba posiada jeden samochód, bardziej maksymalizuje użyteczność niż świat, w którym jedna osoba ma wszystkie pojazdy dla siebie. Kolejną racją za promowaniem równości jest fakt, że nierówności powodują, przynajmniej w wykształconych społeczeństwach, zazdrość, złość czy nienawiść, których nieużyteczność jest oczywista. Według Hare'a, biorąc pod uwagę powyższe dwie cechy, utylitarysta nie musi obawiać się zarzutu, iż jego teoria będzie pochwałać rażące nierówności społeczne³⁷. Jak pisał,

gdybym musiał rozstrzygnąć spór między egalitarystami i ich przeciwnikami, dla typowego zachodniego społeczeństwa przemysłowego wybrałbym umiarkowany egalitaryzm [...]. Jestem przekonany, że faktem jest, iż w systemie dopuszczającym niewolnictwo żadne społeczeństwo się nie wzbogaci i że wszystkie społeczeństwa zdolne funkcjonować w systemie demokratycznym (istotna kwalifikacja) usłyszą dobrą radę, by w imię własnych interesów taki system przyjąć. Przez moje sądy przemawiają jednak przekonania na temat możliwych stanów rzeczy. Gdyby się okazało, że są one fałszywe, wówczas za sprawą tego samego poglądu filozoficznego dotyczącego wywodu moralnego zostałbym rzecznikiem niewolnictwa i tyranii³⁸.

§ 7.3 Stabilność dwóch poziomów myślenia moralnego

Głównym narzędziem, którego używał Hare, udzielając odpowiedzi na przedstawione powyżej wątpliwości, jest dychotomia dwóch płaszczyzn myślenia moralnego: krytycznej i intuicyjnej. Jeśli więc rozgraniczenie, powszechne nauczanie czy praktyczne stosowanie dwóch poziomów myślenia moralnego okazałoby się problematyczne, obrońca utylitarystyki Hare'a byłby w dużych tarapatkach. Próbę podważenia tego podziału przedstawił Bernard Williams. Według niego, człowiek nie jest w stanie spójnie pogodzić ze sobą dwóch tak różnych metod myślenia moralnego³⁹, co zakładał Hare. Myślenie krytyczne wymaga od nas bowiem postawy podejrzliwej, sceptycznej

³⁶ Zob. tenże, *Myślenie...*, dz. cyt., s. 205.

³⁷ Por. tenże, *Ethical Theory and Utilitarianism*, dz. cyt., s. 216.

³⁸ Tenże, *Myślenie...*, dz. cyt., s. 207-208.

³⁹ Zob. B. Williams, *The Structure...*, dz. cyt., s. 189-190.

wobec dotychczasowych standardów, często nonkonformistycznej, indywidualnego traktowania każdej sytuacji i gotowości do śmiałej rewizji zastanych obyczajów moralnych bez negatywnych odczuć. Myślenie intuicyjne natomiast opiera się na całkowicie odmiennych dyspozycjach psychicznych. Robol jest postacią pełną wyrzutów sumienia w przypadku złamania zasad, bezkrytycznej postawy zaufania w stosunku do tradycyjnych reguł. W jaki sposób tych sprzecznych postaw spójnie nauczać, w jaki sposób je zinternalizować? Zdaniem Williamsa, nie można być zarazem krytycznym utilitarystą czynu oraz bezkrytycznym intuicjonistą, który złamanie domniemanych zasad moralnych okupuje poważnymi wyrzutami sumienia.

Trudność polega na tym, generalnie rzecz ujmując, iż nie można myśleć na poziomie intuicyjnym czy codziennym, tak jak teoria tego wymaga, będąc zarazem w pełni świadomym tego, co się czyni: w szczególności w pełni rozumiejąc w terminach samej teorii co czynimy myśląc intuicyjnie⁴⁰.

Jednak zdaniem Hare'a powyższy zarzut ma charakter burzy w szklance wody⁴¹. Polega on na podaniu w wątpliwość tego, iż możemy patrzeć na etykę z dwóch odrębnych perspektyw: z zewnątrz (myślenie krytyczne) i od wewnątrz (myślenie intuicyjne). Czy nie jest jednak tak, że wiele osób (większość filozofów) potrafi być w pełni zaangażowana w akceptowane zasady etyczne, metodologiczne, religijne czy polityczne i jednocześnie zastanawiać się nad ich uzasadnieniem, prawdziwością czy racjonalnością, zmieniając perspektywę w zależności od potrzeb i okoliczności? Gdybyśmy mieli zaakceptować sposób rozumowania Williamsa, należałoby przystać na to, że nikt, kto raz na jakiś czas krytycznie rozmyśla na temat swojego życia nie mógłby być w pełni w nie zaangażowany.

§ 7.4 Cnoty a utilitaryzm

Człowiek posługuje się najczęściej myśleniem intuicyjnym. Istnieją jednak sytuacje, w których korzysta z myślenia krytycznego⁴². Zdaniem Hare'a, nie jest bowiem tak, iż myślenie krytyczne jest skutecznym narzędziem jedynie w rękach archanioła. Choć przedstawiając swój model utilitaryzmu reguł-preferencji, Hare opisywał go jako system

wzorowy, skonstruowany dla idealnego nakazodawcy, również my możemy i mamy w nim udział. Potrafimy myśleć krytycznie, choć wymaga to od nas dużego wysiłku i zdolności. Chociaż istnieją konflikty, których rozwiązaniu sprostać może jedynie idealny nakazodawca⁴³, w wielu sytuacjach myślenie krytyczne stosujemy z powodzeniem. Wytworzyliśmy bowiem narzędzia, które zbliżają nas do archanioła.

Wbrew kolejnemu rozpowszechnionemu zarzutowi, że utilitaryzm, skupiając się na zasadach, nie potrafi uzasadnić wagi cnót i wad w zrozumieniu moralności, wyrabianie pożądanych cech charakteru zdaniem Hare'a ma fundamentalne znaczenie dla codziennego namysłu moralnego oraz tego, aby czynić dobro.

Żeby myśleć krytycznie, człowiek powinien mieć wiele cech, które nazwać można cnotami intelektualnymi. Najważniejszą z nich, tak jak dla Arystotelesa, jest **rozsądek** (*phronesis*) – umiejętność posługiwania się racjonalnością praktyczną. Sprawność tę można rozwijać poprzez **doświadczenie**⁴⁴. Pomoże nam ono rozstrzygnąć, kiedy myśleć krytycznie, a kiedy stosować zasady domniemane. Osoba posiadająca bogate doświadczenie będzie miała pewne metodologiczne zasady domniemane, które podpowiedzą jej, kiedy uruchomić myślenie krytyczne.

Inną ważną cnotą, niezbędną do zdobycia informacji koniecznych do przeprowadzenia myślenia krytycznego, jest **wyobraźnia**⁴⁵ i **umiejętność przewidywania**. Nie posiadamy bowiem, tak jak archanioł, wiedzy dotyczącej skutków danego działania czy wprowadzenia reguły. Z tego względu człowiek, jeśli chce dokonać wyboru działania posługując się myśleniem krytycznym, powinien ocenić prawdopodobieństwo wystąpienia jego ewentualnych skutków i wziąć je pod uwagę⁴⁶.

Gdy nie posiadamy pełnych informacji, najlepszą rzeczą, jaką możemy zrobić, to działać na podstawie informacji, które posiadamy, dokonując najbardziej racjonalnej decyzji w świetle oszacowanego prawdopodobieństwa⁴⁷.

Inną ważną cechą jest **wrażliwość**⁴⁸. Dzięki niej w procesie uniwersalizacji możemy skutecznie utożsamić się z innymi ludźmi. Bez wy-

⁴⁰ Tenże, s. 192.

⁴¹ Zob. R. M. Hare, *Comments...*, dz. cyt., s. 289.

⁴² Zob. tenże, *Myślenie moralne*, dz. cyt., s. 61-62, 70.

⁴³ Tamże, s. 142.

⁴⁴ Zob. tamże, s. 70.

⁴⁵ Zob. tenże, *Freedom...*, dz. cyt., s. 95.

⁴⁶ Zob. tenże, *A Utilitarian Approach...*, dz. cyt., s. 155.

⁴⁷ Tamże.

⁴⁸ Zob. tenże, *Freedom...*, dz. cyt., s. 181-183.

obraźni, wrażliwości i empatii nie moglibyśmy np. zrozumieć, jak to jest być nieuleczalnie chorym i cierpiącym, co może być konieczne, choćby do opracowania zasad etyki lekarskiej. W codziennej praktyce moralnej niezbędne są również cnoty, które Hare nazwał **instrumentalnymi**: odwaga, opanowanie, wytrwałość, umiarkowanie. Bez nich nie można być nawet skutecznym złodziejem⁴⁹. Powszechnie akceptowane zasady moralne stwarzają więc ludzie wielcy, tacy, którzy są doświadczeni, wrażliwi i mają wyobraźnię. Ci, którzy lękają się myślenia krytycznego, mogą i powinni czerpać z mądrości pokoleń⁵⁰.

Posiadanie cnót ma jednak również fundamentalne znaczenie w codziennej praktyce moralnej. Pamiętać bowiem należy, że dobre działania są zdefiniowane przez to, że dokonuje ich człowiek cnotliwy⁵¹. Człowiek moralnie dobry posiada wiele, jak można byłoby powiedzieć, domniemanych cnót moralnych, które swą treść czerpią ze skorelowanych z nimi zasad domniemanych. **Cnoty domniemane** byłyby więc trwałymi dyspozycjami do zachowania się zgodnie z wyuczonymi zasadami domniemanymi. Jak już mówiliśmy, jedynie istnienie takich zinternalizowanych dyspozycji może gwarantować przestrzeganie wielu ważnych zasad i reguł z płaszczyzny intuicyjnej. Aretologia – dział etyki zajmujący się cnotami – pełni więc istotną funkcję zarówno dla poziomu krytycznego, jak i codziennego działania opartego na namyśle intuicyjnym.

§ 7.5 Utylitaryzm a myślenie intuicyjne

Jak pisaliśmy w rozdziale piątym, myślenie krytyczne jest formą utylitaryzmu. Wraz z wprowadzeniem dwóch poziomów myślenia moralnego Hare dodatkowo skomplikował swoją wersję tego stanowiska: przede wszystkim rozdzielił czyny moralnie słuszne i niesłuszne od dobrych i złych, ufundował codzienne decyzje moralne na intuicjach prostych zasad oraz nadał cnotom bardzo ważną funkcję. Dzięki powyższym modyfikacjom jego teoria broni wielu fenomenów związanych z codzienną praktyką moralną, które były podstawą rozwijania deontologii i intuicjonizmu. Zachował on jednak trzon utylitarystyczny: płaszczyzna intuicyjna nie jest wyłomem i niekon-

sekwencją jego teorii, uzyskała bowiem uzasadnienie dzięki metodzie krytycznej zastosowanej do pewnych trudno podważalnych założeń psychologicznych dotyczących cech człowieka.

Zdaniem Hare'a, jeśli myślenie archanioła w praktyce pokrywa się z umiarkowanym utylitaryzmem czynu; to płaszczyzna intuicyjna z pewnością jest formą utylitaryzmu reguł. To, co różni bowiem umiarkowanego utylitarystę czynu od utylitarysty reguł, to stosunek do zasad: dla pierwszego zasada, np. nie wolno łamać obietnic, jest jedynie praktycznym, użytecznym, indukcyjnym uogólnieniem (*rules of thumb*), które w odpowiednim momencie, tj. jeśli okoliczności ulegną zmianie, należy porzucić; dla drugiego zaś zasady mają charakter moralny i z ich złamaniem słusznie związane jest poczucie winy, chęć zadośćuczynienia i społeczna nagana⁵².

W celu dokonania wyboru zasad domniemanych i związanych z nimi dyspozycji do zachowania⁵³ powinniśmy mieć na uwadze globalny efekt – maksymalizację spełnienia preferencji, która będzie rezultatem kultuwowania czy zaadoptowania dyspozycji i testowanych zasad. Dlatego też etykę płaszczyzny intuicyjnej nazwać można **utilitaryzmem reguł kultuwowanych**. W praktyce oznacza to, iż utylitarysta, którego celem jest zalecenie systemu domniemanych zasad i cnót dla danego społeczeństwa, powinien wziąć pod uwagę co najmniej takie czynniki, jak: (1) bezpośrednie koszty i zalety wprowadzania dyspozycji i reguł; (2) skuteczność wprowadzonych dyspozycji, które powinny powodować pożądane działania; (3) koszty zabezpieczenia, aby dyspozycje czy reguły były przestrzegane; (4) zasięg zalet i kosztów, które rodzą się z przestrzegania reguł; (5) prawdopodobieństwo wszystkich tych konsekwencji; (5) oczekiwania ludzi związane z tym, aby inni przestrzegali danych reguł, które będą miały wpływ na społeczną motywację do działania zgodnego z regułami⁵⁴.

⁵² Zob. W. H. Show, *Contemporary Ethics...*, dz. cyt., s. 159, 164.

⁵³ Zob. § 6, poświęcony cnotom w myśli R. M. Hare'a.

⁵⁴ Zob. R. M. Hare, *Comments on Frankena*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 225-226. Powyższe zależności są odpowiedzią Hare'a na zarzut Williama Frankena, iż Hare w swoich książkach jasno nie przedstawił, w jaki sposób za pomocą myślenia krytycznego dokonujemy wyboru, uzasadniamy czy testujemy zasady domniemane. Zob. W. K. Frankena, *Hare on the Levels of Moral Thinking*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 48.

⁴⁹ Zob. tenże, *Myślenie...*, dz. cyt., s. 240.

⁵⁰ Zob. tamże, s. 64.

⁵¹ Zob. § 7.1.1.

W związku z podziałem sposobu myślenia moralnego na dwie płaszczyzny William Frankena podniósł jednak wątpliwość, czy uda nam się zaklasyfikować normatywną działalność człowieka, który najczęściej myśli jak deontolog-intuicjonista-robot, w szczególnych przypadkach natomiast powinien prowadzić rozważania jak archanioł-utylitarysta czynu⁵⁵. Przykładem egzemplifikującym to wewnętrzne rozbieżności jest również definiowanie pojęć czynu moralnie słusznego i czynu moralnie dobrego: pierwsze Hare wyznacza poprzez metodę myślenia krytycznego, drugie natomiast w odniesieniu do cnót, które swą treść czerpią z zasad interpretowanych w duchu utilitaryzmu reguł. Zdaniem Hare'a, ścisła dychotomia dwóch rodzajów utilitaryzmu (czynu i reguł) nie jest jednak adekwatnym narzędziem do opisu jego dwupoziomowego utilitaryzmu preferencji⁵⁶. Uważał on, że jest

zarazem utylitarystą czynu jak i reguł: dyspozycje do działania z płaszczyzny intuicyjnej są bowiem bliskie temu, co w literaturze nazywa się „regułami”⁵⁷,

natomiast

to, co w istocie selekcjonujemy, to czyny związane z przestrzeganiem i kultywowaniem różnych dyspozycji. Oceniając owe działania (co musi być dokonane globalnie w świetle konsekwencji całej sieci czynów), musimy używać myślenia krytycznego⁵⁸.

Według niego, choć prezentowane stanowisko nie pokrywa się z utilitaryzmem reguł w rozumieniu R. Brandta, uzasadnia ono postępowanie zgodne z regułami, wskazując na użyteczność serii czynów, które wynikają z ich przestrzegania i kultywowania. Z tego względu unika ona klasycznych problemów związanych z prostą wersją utilitaryzmu czynu⁵⁹.

⁵⁵ Zob. W. K. Frankena, *Hare...*, dz. cyt., s. 43-56.

⁵⁶ Zob. R. M. Hare, *Comments on Harsanyi*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 242.

⁵⁷ R. M. Hare, *Comments on Frankena*, w: D. Seanor, N. Fotion (ed.), *Hare and Critics...*, dz. cyt., s. 226.

⁵⁸ Tamże, s. 227.

⁵⁹ Zob. tamże.

Zakończenie

Celem niniejszej pracy było przedstawienie, analiza oraz krytyczna ocena filozofii moralnej Richarda Marvina Hare'a. W tym miejscu postaramy się podsumować to, co zostało powiedziane, oraz wysnuć końcowe refleksje.

Filozofia moralności Richarda Hare'a wyrasta z kilku źródeł: krytyki naturalizmu dokonanej przez G. E. Moore'a (§ 2.1); emotywiizmu (§ 2.2); szkoły analizy języka potocznego (§ 2.3); oraz filozofii I. Kanta i utylitarystów (§ 2.4). Głównym założeniem Hare'a było przeświadczenie, że za pomocą analizy lingwistycznej można dojść do fundamentalnych dla etyki wniosków. Aby odpowiedzieć na nurtujące ludzi pytania, w pierwszym rzędzie należy poznać język, w którym są one formułowane (§ 2.3). Metaetyka jako analiza znaczeń oraz logiki języka moralnego nie ma więc jedynie charakteru opisowego, lecz pośrednio jest dziedziną normatywną. Często bowiem z powodu błędnych założeń na temat charakteru języka oraz związanej z nim metodologii etyki nie potrafimy poradzić sobie z praktycznymi problemami moralnymi, które mogą nam się jawić jako nierozwiązywalne dylematy (§ 3.1). Z tego względu należy najpierw wyleczyć filozofów moralności z błędnych mniemań i metaetycznych opinii (§ 2.3).

Z badań nad językiem potocznym wynika, iż język moralny służy do przekazywania w ekonomiczny sposób zaleceń, rad i nakazów (§ 3.1). Dobra teoria metaetyczna powinna brać to pod uwagę. Spełnić musi ona również szereg wymagań, które nazwaliśmy warunkami pragmatycznymi (§ 3.4). Jednak większość współczesnych koncepcji metaetycznych (§ 3.2) zawodzi w tym względzie: emoty-

wizm prowadzi do irracjonalizmu oraz wielu innych problemów (§ 3.3.1); deskrytywizm popada w relatywizm i niemożność wyjaśnienia fenomenu sporu moralnego (§ 3.3.2); teorie realistyczne nie są spójne z szeroko akceptowanymi założeniami filozoficznymi i naukowymi oraz są niepraktyczne (§ 3.3.3); wreszcie żadna licząca się wersja deskrytywizmu nie potrafi wyjaśnić koniecznej relacji między motywacją a powinnością (§ 3.3.4). Dlatego rację bytu ma jedynie stanowisko racjonalnego nondeskrytywizmu. Jego realizacją, według Hare'a, jest uniwersalny preskrytywizm.

Aby rzetelnie ocenić teorię metaetyczną, nie można brać pod uwagę jedynie kryteriów przedstawionych przez Hare'a. Należy zastanowić się również nad tymi, które były postulowane przez jego krytyków (§ 3.4). Dobra koncepcja metaetyczna, oprócz realizacji praktycznego celu etyki, powinna wyjaśniać również szereg fenomenów związanych z moralnością oraz być spójna z powszechnie uznawanymi założeniami naukowymi i filozoficznymi. Biorąc to pod uwagę, stwierdzić należy, że Hare, pomimo iż próbował zdać sprawę z faktycznego funkcjonowania języka moralnego, był zwolennikiem rewizjonistycznej odmiany teorii metaetycznej (§ 4.1). Z podejrzliwością patrzył bowiem na niektóre intuicje lingwistyczne, *de facto* nadając im charakter epifenomenalny. Starał się dowieść, że sądy moralne, choć w swojej strukturze powierzchniowej podobne są do zdań opisowych i z tego powodu wydają się być wyrazem prawdziwych lub fałszywych przekonań, różnią się od nich. Chociaż jawią się jako stwierdzenia, jest to jedynie ich własność powierzchniowa, w swej strukturze głębokiej bowiem posiadają cechy odmienne. Ponieważ także imperatywy mają niektóre z tych cech, aby poznać sądy moralne, należy studiować imperatywy (§ 4.1).

Oceny nie są jednak nakazami (§ 4.1.2, § 4.3), lecz formą preskrypcji (§ 4.1). Mają charakter nadrzędnych, uniwersalizowalnych zaleceń, co najlepiej można wyjaśnić w języku teorii czynności mowy (§ 4.1.2). Choć takie ich definiowanie rodzi pewne problemy (§ 4.2), z których najpoważniejszy został wyrażony w argumencie Fregego-Geacha (§ 4.7), Hare starał się je przezwyciężyć (§ 4.8). Znaczenie podstawowych słów moralnych takich, jak „powinno się” czy „dobry”, ma charakter czysto formalny (§ 4.10). Są one operatorami logiki moralnej gry językowej – o ich znaczeniach jesteśmy w stanie dowiedzieć się wszystkiego, jeśli wiemy, jakie wnioski możemy wy-

ciągnąć ze zdań, które je zawierają (§ 4.10). Na znaczenie pojęć i sądów moralnych nie mają wpływu intencje nadawcy, perlokucyjne czynności językowe (§ 4.9), lecz konwencjonalna illokucyjna czynność zalecania, ich lokucja oraz inne formalne własności lingwistyczne. Znaczenie słów moralnych nie sprowadza się również do kryteriów ich aplikacji (§ 4.6), warunki te bowiem są zmienne – zależą od czasu oraz miejsca.

Hare, aby udowodnić, że jego teoria metaetyczna umożliwia racjonalną argumentację w obrębie etyki, a sądy moralne mogą rządzić się prawami logiki, starał się wykazać, iż nawet najpospolitsza i najbardziej problematyczna z punktu widzenia logiki klasycznej forma zaleceń, jaką są imperatywy, podlega jej prawom (§ 4.11). Nie istnieje jednak specyficzna logika imperatywów, rządzą się one bowiem tymi samymi regułami co zdania oznajmujące. Logikę tę można opisać w terminach niezależnych od pojęcia prawdy: jest ona zbiorem zasad kierujących poprawnymi sposobami wnioskowania bez względu na to, czy przesłanki są prawdziwe, fałszywe, czy nie posiadają tych wartości (§ 4.11). Rodzaj aktu mowy, który stanowi podstawę wnioskowania, jest jednak ważny. Żaden wniosek imperatywny nie może być poprawnie wyprowadzony ze zbioru przesłanek, które nie zawierają przynajmniej jednego imperatywu (§ 4.12). Niewzięcie tego pod uwagę jest jedną z przyczyn nieadekwatności projektu naturalistycznego.

Z zaleceniami związana jest bardzo ważna cecha: szczerą ich akceptacją motywuje do działania. Z tego względu sądy moralne podlegają prawu internalizmu motywacyjnego (§ 4.13): szczerą akceptacją sądu: „Powinno się wykonać x” pociąga za sobą konieczność posiadania motywu do wykonania x. Zasada ta ma jednak pewne ograniczenia: są nimi problem słabości woli (§ 4.13.1) oraz amoralizm (§ 4.13.2). Ich analiza może przekonać nas, iż Hare nie był zwolennikiem Hume'owskiego modelu motywacyjnego. W kwestii psychologii moralnej bliższy był mu bowiem Kant i Arystoteles (§ 4.13.1). Zalecenia nie muszą być ekspresją jedynie naszych życzeń czy uczuć, ale mogą być również wyrazem racjonalnej woli. Analiza problemu akrazji skłania także do poglądu, że Hare nie był przedstawicielem mocnego internalizmu motywacyjnego. Choć akceptował konieczność relacji pojęciowej między motywem a obowiązkiem, nie postulował koniecznego związku obowiązku i działania (§ 4.13.1). Uważał natomiast, iż sądy moralne są

nadrzędne: jeśli ktoś szczerze używa języka moralnego, to wszystkie zasady, które są w konflikcie z prawami moralnymi (np. zasady estetyczne, egoistyczny interes własny, wybory ze względów politycznych czy ekonomicznych) musi traktować jako podrzędne wobec zasad moralnych i w razie konfliktu je uchylić (§ 4.14).

Sądy moralne, choć różnią się od sądów opisowych, pod pewnym względem podobne są jednak do przekonań. Nie mogą być, w przeciwieństwie do imperatywów, wypowiedziane arbitralnie, lecz muszą opierać się na racjach i uzasadnieniach (§ 5.1). Są one także uniwersalizowalne (można uczynić je uniwersalnymi) (§ 5.2). Uniwersalny sąd moralny to taki akt mowy, który jest zaleceniem bez stałych indywidualnych (np. nazw osób, miejsc, czasu) oraz został wyrażony jedynie za pomocą predykatów odnoszących się do własności uniwersalnych.

Uniwersalność jest związana z prawem superweniencji. Hare nie akceptował jednak superweniencji metafizycznej, twierdził jedynie, że jeśli dwa przedmioty czy czyny są rozpoznawane przez dwie osoby jako posiadające te same zbiory cech moralnie istotnych, to w odniesieniu do nich muszą one zastosować te same predykaty moralne (§ 5.2). Superweniencja taka nie jest mocnym prawem metafizycznym, lecz regułą pewnych aktów mowy zrelatywizowanych do danych osób. Z tego względu można powątpiewać, czy zabezpieczy ona obiektywność sądów moralnych. Postawić można zarzut, iż niezgodna na konieczną, analityczną lub metafizyczną relację między pewnymi nierelatywnymi predykatami (własnościami) naturalnymi a predykatami (własnościami) moralnymi będzie prowadzić do relatywizmu (§ 5.2). Aby uniknąć takiej konkluzji, Hare musiał postulować możliwość wyznaczenia poprzez wyidealizowaną procedurę absolutnych, niezwiązanych z jednostkowym punktem widzenia, istotnych cech naturalnych, na których powinny opierać się oceny. Rozpoznać je można dzięki teorii relewancji (§ 6.2).

Choć w dziedzinie metaetyki Hare był zwolennikiem teorii rewizjonistycznej, nie jest jednak prawdą, iż odrzucał możliwość przypisywania wartości logicznych prawdy lub fałszu sądom moralnym. Nie akceptował jednak korespondencyjnej teorii prawdy (§ 5.3). W pierwszym okresie był deflacionistą, później natomiast popierał minimalną teorię prawdy C. Wrighta. Z pełną otwartością głosił, że sądy moralne mogą być prawdziwe lub fałszywe (§ 5.3). Myślenie

krytyczne przeprowadzone przez archanioła można traktować jako metodę dojścia do superstwierdzalności C. Wrighta: moralnie prawdziwy byłby sąd, który jest moralnie uzasadniony i takim pozostaje bez względu na to, jak szczegółowe i doskonałe podjęlibyśmy w tej kwestii dodatkowe rozważania (§ 5.3). Obiektywność sądów etycznych jest zatem zagwarantowana nie poprzez istnienie moralnych faktów, lecz dzięki procesowi rozumowania moralnego, który przeprowadzony przez archanioła, doprowadza do uznania niezmiennej i jednej oceny danej sytuacji, którą nazywamy prawdziwą.

Drugim, po rozwiązaniach metaetycznych, istotnym elementem systemu etycznego Hare'a jest jego normatywna teoria etyczna, którą podzielił na dwie płaszczyzny: intuicyjną i krytyczną. Dychotomia ta jest związana z dwoma sposobami myślenia moralnego i dwoma rodzajami reguł (§ 6.1). Celem myślenia krytycznego jest dostarczenie idealnej, wzorcowej metody rozwiązywania trudnych moralnych problemów (§ 6), celem myślenia intuicyjnego zaś podanie sposobu udzielenia odpowiedzi na wymagania etyki dnia powszedniego (§ 7). Wzorowym przedstawicielem myślenia krytycznego jest idealny nakazodawca – archanioł, bohaterem myślenia intuicyjnego natomiast – robot.

Ponieważ myślenie krytyczne powinno polegać na uniwersalizacji nakazów na podstawie istotnych własności ocenianej sytuacji (§ 6.3), można wyróżnić różne jego etapy ze względu na podział własności relewantnych. Pierwszy (§ 6.3.1) oraz drugi (§ 6.3.2) etap myślenia krytycznego, według Hare'a, mają charakter formalny i wynikają z bronionych tez metaetycznych. Polegają one na związaniu kwantyfikatorem ogólnym istotnych formalnych cech sytuacji etycznej. Wyprowadzając jakiś nakaz, musimy założyć, że jest on obowiązujący we wszystkich przypadkach, które są identyczne pod względem moralnie istotnych własności oraz bez względu na rolę, jaką mielibyśmy w niej przyjąć (lub która byłaby nam w niej narzucona). Trzeci etap (§ 6.3.3.) nie jest jednak czysto formalny, gdyż zakłada *zasadę pełnego utożsamienia się z innym*, racjonalność podmiotów działających oraz wymaga posiadania wiedzy na temat istotnych materialnych własności dotyczących ocenianej sytuacji. Z tego względu etyka normatywna Hare'a nie wynika jedynie z tez metaetycznych. Metaetyka stanowi dla niej fundament jedynie w sensie kantowskim, nie zaś kartezjańskim (§ 6.2). Co więcej, również pre-

skryptywność i uniwersalność języka moralnego udowodnić można jedynie nie wprost. Aby to uczynić, należy wykazać, iż uniwersalnego preskryptywizmu nie można skutecznie obalić, że dobrze spełnia on swą praktyczną funkcję oraz realizuje listę wymagań dla dobrej teorii metaetycznej (§ 6.2). Aby móc utrzymać myślenie krytyczne jako teorię spójną, Hare musiał bronić również nondeskrytywistycznej teorii tożsamości osobowej (§ 6.4). Według tej koncepcji, „ja” nie jest słowem do końca opisowym, lecz po części nakazowym: utożsamiając się z jakąś osobą, faktycznie albo hipotetycznie identyfikujemy się z jej nakazami.

Pomimo istotnych różnic między klasycznym utylityzmem a myśleniem krytycznym zaproponowaną przez Hare’a metodę myślenia krytycznego należy nazwać utylityzmem (§ 6.5). Ma ona formę wyrafinowanego bądź umiarkowanego utylityzmu czynu, w którym należy maksymalizować spełnienie racjonalnych preferencji. Wraz z wprowadzeniem dwóch płaszczyzn myślenia moralnego utylityzm Hare’a uległ dalszym modyfikacjom (§ 7.5). Dla robota, bohatera myślenia intuicyjnego, namysł moralny przybiera bowiem formę globalnego utylityzmu kulturowanych reguł. Człowiek natomiast, połączenie archanioła i robota, musi myśleć raz jak pierwszy, innym razem zaś jak drugi. Ta dwuaspektowość podejścia do problemów moralnych spotkała się z krytyką, którą Hare starał się odeprzeć (§ 7.3).

Choć utylityzm klasyczny boryka się z wieloma problemami, według Hare’a broniony przez niego sposób myślenia moralnego umożliwia ich przewyciężenie. Na zarzut, że utylityzm jest zbyt wymagający (§ 7.1), odpowiedział on rozróżnieniem definicji czynu słusznego, czynu dobrego, kryterium wyboru właściwego działania oraz warunków bycia winnym jego niespełnienia (§ 7.1.1). Uzasadniał i bronił również podziału działań na obowiązkowe i nadobowiązkowe (§ 7.1.2) oraz zalecania przedkładania interesów bliskich nad osoby obce (§ 7.1.3). Starał się także wyjaśnić, w jaki sposób w ramach jego wersji utylityzmu można uzasadnić pojęcie sprawiedliwego podziału dóbr, które musi zakładać pewną dozę egalitaryzmu (§ 7.2). Wskazywał także na wagę cnót, które stanowią istotę człowieka moralnie dobrego oraz pozwalają nam zbliżyć się do sposobu myślenia archanioła (§ 7.4).

* * *

W interpretacji przedstawionej i uzasadnianej na kartach tej książki niektóre z cech metaetyki oraz etyki normatywnej Hare’a różnią się od często mu przypisywanych. Po pierwsze, Hare nie redukował sądów moralnych do imperatywów czy zaleceń, lecz analizował znaczenie pojęć i sądów w odniesieniu do formalnych reguł ich użycia – pojęciowych praw, które są wynikiem ich stosowania (§ 4.3). Po drugie, nie był nonkognitywistą, lecz nondeskrytywistą akceptującym kognitywizm (prawdziwość/fałszywość sądów moralnych oraz istnienie moralnej wiedzy) (§ 3.2, § 5.3). Po trzecie, nie był on również zwolennikiem mocnego internalizmu ani Hume’owskiego modelu motywacyjnego (§ 4.13). Po czwarte, nie był relatywistą, lecz absolutystą (metodę myślenia krytycznego należy traktować jako epistemologiczny proces dojścia do prawdziwości sądów moralnych – superstwierdzalności) (§ 5.3). Cechy te powodują, że nie powinno traktować się go jako emotywisty lub postemotywisty, lecz raczej wymieniać jednym tchem ze współczesnymi przedstawicielami wyrafinowanego nondeskrytywizmu (S. Blackburnem, A. Gibbardem, M. Timmonsem), którego był prekursorem. Wbrew swoim krytykom nie próbował on również wywieść etyki normatywnej jedynie z logiki języka (§ 6.2). Kolejną cechą odróżniającą rozwijaną tu interpretację od standardowego wykładu etyki Hare’a jest zwrócenie uwagi na to, iż bronił on utylityzmu racjonalnych preferencji (§ 6.3.4), w którym ważny był globalny skutek wynikający z kulturowania reguł i cnót (§ 7). Potrafił on również skutecznie wyjaśnić wiele wątpliwości nagromadzonych wokół utylityzmu (§ 7.1, § 7.2, § 7.3).

W poszczególnych paragrafach niniejszej książki, ustosunkowując się do kolejnych argumentów, staraliśmy się podać odpowiedź na pytanie, czy współcześnie, biorąc pod uwagę liczne argumenty krytyczne skierowane pod adresem koncepcji metaetycznej i normatywnej Hare’a, warto być zwolennikiem tych teorii. W tym miejscu zatem zbierzemy jedynie najważniejsze z tych opinii oraz dodamy kilka nowych uwag, które, miejmy nadzieję, rzucają nieco światła na tak postawione pytanie.

Wielu filozofom stanowisko metaetycznego deskrytywizmu jawi się jako najbardziej godne akceptacji. Nie bez powodu nondeskrytywizm jest nazywany rewizjonizmem. Takie cechy sądów etycznych, jak: stwierdzanie faktów, bycie prawdziwymi/fałszywymi, posiadanie charakteru poznawczego, konieczność istnienia uzasadniającej

je racji, bycie przesłankami lub wnioskiem rozumowań, stanowią argument przeciwko każdemu pogładowi metaetycznemu, który nie bierze ich pod uwagę (§ 3.4). Deskrytywista nie ma problemu z ich wyjaśnieniem, wynika ono bowiem z poczynionych przez niego założeń. Z drugiej jednak strony, nondeskrytywizm, w tym również uniwersalny preskrytywizm, rozwiązuje problemy, z którymi kłopoty mają jego adwersarze. Teoria Hare'a ma bowiem szereg zalet. Niektóre dzieli z emotywizmem, inne stanowią jej cechą charakterystyczną. Do pierwszych należy możliwość wyjaśnienia w prosty sposób: poczucia odrębności języka moralnego i naukowego (§ 4.1); fenomenu „otwartości” definicji terminów moralnych w terminach opisowych (§ 3.3.4); motywacji do czynienia tego, co uważamy za słuszne (§ 4.1); problemu przejścia od „jest” do „powinien” (§ 4.12). Dodatkowymi zaletami uniwersalnego preskrytywizmu są: założenie istnienia w języku potocznym logiki deontycznej i próba jej odтворzenia (§ 4.11); akceptacja i wyjaśnienie racjonalności sporu moralnego oraz potrzeby posiadania przez sądy moralne racji uzasadniającej (§ 5.1); akceptacja i wyjaśnienie faktu prawdziwości bądź fałszywości sądów moralnych (§ 5.3); proste wyjaśnienie fenomenu sporu moralnego (§ 4.6), zjawiska reformy moralnej i sensowności (rozumiałości) użycia podstawowych pojęć moralnych do oceny nieznanych wcześniej czynów, sytuacji i przedmiotów (§ 4.6); wskazywanie etyki normatywnej, która może zapewniać absolutność i obiektywność ocen moralnych – metody myślenia krytycznego podlegającej zasadom utilitaryzmu reguł-preferencji (§ 6.2); proste wyjaśnienie globalnej superwencji predykatów moralnych na predykatkach naturalnych bez powodowania, jak w przypadku realizmu, metafizycznych problemów i zawilości (§ 2.2, § 5.2). Poza tym uniwersalny preskrytywizm dobrze radzi sobie ze spełnieniem większości warunków, które za M. Timmonsem nazwaliśmy kryteriami koherencji oraz fenomenologii (§ 3.4). Ze zbiorem pierwszych z nich nie ma on problemu, ponieważ w szerokim sensie tego słowa, jest to projekt naturalistyczny: Hare zgodziłby się, iż ontologicznie pierwotne są wszystkie te byty (własności, fakty itd.), które są przedmiotem nauki (KH1), każdy dyskurs pozanaukowy musi być dostosowany do naukowego (naturalistycznego) obrazu świata (KH2) oraz wiedza i uzasadnienie związane z dyskursem pozanaukowym muszą być zgodne z naturalistycznym sposobem wyjaśniania (KH3) (§ 3.3.3). W ra-

mach jego teorii nie ma również problemu z wyjaśnieniem fenomenu głębokich i autentycznych różnic poglądów moralnych (KF3) (§ 4.6); motywowania przez sądy moralne do działania (KF4) (§ 4.13); bazowania sądów moralnych na autorytecie rozumu (KF5) (§ 6.2) oraz ich nadrzędności nad innymi praktykami oceniania (KF6) (§4.14). Uniwersalny preskrytywizm, w przeciwieństwie do emotywizmu (§ 3.3.1), jednoznacznie opisuje także cechy specyficzne dla języka moralnego (§ 4.1), wyjaśnia, dlaczego wydając oceny, nie musimy przyjmować w stosunku do nich żadnych postaw (§ 4.3, § 4.9), oraz opisuje relację występującą pomiędzy elementem deskryptywnym a pozadeskryptywnym w znaczeniu sądów etycznych (§ 4.6, 6.2).

Hare próbował z mniejszym lub większym powodzeniem poważnie traktować przedstawione przez nas fenomeny moralności (§ 3.4). Starał się pogodzić mocny internalizm pojęciowy ze słabością woli oraz amoralizmem (§ 4.13); nondeskryptywny charakter sądów etycznych z posiadaniem przez nie wartości logicznych (§ 5.3); definiowanie znaczeń poprzez moc illokucyjną z logicznością sądów moralnych (§ 4.5) oraz ich znaczeniem w kontekstach złożonych (argumentem Fregego–Geacha) (§ 4.8). Starał się również odpowiedzieć na zarzut popadania w relatywizm koncepcji, w której zakłada się rozdzielenie kryterium użycia słów i zdań od ich znaczenia (§ 6.2). Można jednak argumentować, że uniwersalny preskrytywizm nie radzi sobie z wyjaśnieniem takich fenomenów, jak KF1 (sądy moralne są standardowo wypowiedzane jako stwierdzenia w trybie oznajmującym i wydają się wyrażać autentyczne przekonania) oraz KF2 (sądy moralne mogą być prawdziwe lub fałszywe i faktycznie wiele z nich jest prawdziwych) (§ 3.4). Choć Hare z biegiem czasu przypisywał im coraz większą rolę, starając się zawrzeć zwłaszcza drugi z nich w ramach uniwersalnego preskrytywizmu, można twierdzić, że rezultat tych zabiegów nie jest satysfakcjonujący. Nie każdy jest przekonany do deflacyjnej czy minimalnej teorii prawdy lub minimalnych stwierdzeń. Czy nie jest tak, że posługiwanie się przez deskrytywizm autentycznym, mocnym pojęciem prawdy i stwierdzenia stanowi o jego przewadze? Co więcej, minimalizm, nie jest autorstwa Hare'a, choć stanowi spójne i twórcze rozwinięcie jego koncepcji. Przedstawione przez niego autorskie rozwiązanie *problemu uwikłania*, choć stara się ochronić uniwersalny

preskrytywizm przed argumentem Fregego–Geacha, wymaga natomiast dalszego, szczegółowego opracowania (§ 4.7, § 4.8).

Powody odrzucenia przez Hare'a deskrytywizmu także w większości są już nieaktualne: problem z *testem otwartego pytania* oraz niemożność wyjaśnienia istotnej dla moralności relacji między motywacją a powinnością nie stanowi obecnie dla jego przedstawicieli nierozwiązywalnego problemu. Z pierwszym radzi sobie bowiem np. realizm z Cornell (§ 3.3.2) czy redukcyjny, lecz rewizjonistyczny naturalizm P. Railtona¹, drugi natomiast stara się rozwiązać np. deskryptywny internalizm T. Nagela² i J. McDowella³, podważający założenie Hume'a, iż przekonania są motywacyjnie bierno, bądź, z drugiej strony, internalistyczny naturalizm M. Smitha⁴ lub P. Railtona⁵, który akceptuje tezę o motywacyjnej niemocy przekonań. Hare suponował bowiem, że stany poznawcze, które polegają na uchwytowaniu prawdy, są niezdolne do motywowania (§ 4.13.2.). Z tego też powodu uważał, że to, co M. Smith traktuje jako rozwiązywalny moralny problem⁶, stanowi dylemat nie do pogodzenia: albo przyjmujemy istnienie koniecznej relacji między motywem a powinnością, albo między oceną a faktem; albo uważamy, iż sądy moralne mają charakter nakazowy, albo wyrażają prawdziwe bądź fałszywe opisy (§ 4.13.1). Również pozytywistyczne założenie o radykalnej różnicy faktów i wartości, opisowości i normatywności, na którym opiera się jego nondeskrytywizm, podane zostało w wątpliwość⁷. Biorąc to wszystko pod uwagę, można zasadnie twierdzić, że nie ma obecnie powodu, aby akceptować nondeskrytywizm.

Jednak niezgoda Hare'a na deskrytywizm opierała się przede wszystkim na innym argumencie: jego niepraktyczności sprowadzającej się do *reductio ad absurdum* tego projektu (§ 3.3.2). Według niego, deskrytywiści zaprzepaszczają praktyczny wymiar etyki: uniemożliwiają odkrycie spójnej z ich poglądami etyki normatyw-

nej, która pozwalałaby rozwiązywać konflikty moralne dzięki rozmowi. W obliczu pytania: „Co powinniśmy w danej sytuacji czynić?” możliwe są bowiem co najmniej cztery sposoby udzielenia odpowiedzi: odwołanie się do podzielanych intuicji moralnych; próba odkrycia obiektywnych, realnych moralnych faktów (naturalistycznych bądź nonnaturalistycznych); poszukiwanie apriorycznego, czysto rozumowego dowodu lub uzgadnianie ludzkich postaw, pragnień, preferencji i interesów. Według Hare'a, etyka normatywna o ambicjach praktycznych nie może zostać jednak ufundowana na żadnych materialnych konwencjach moralnych czy językowych (§ 3.3.2), sumieniu, naszych intuicjach lub przekonaniach (§ 3.3.3). Tą drogą nie uda nam się bowiem rozwiązać istniejących sporów moralnych. Wielu z nas co innego rozumie przez szczęście czy dobrobyt; sumienie czy intuicje podają natomiast odpowiedzi sprzeczne. Nie uda nam się również odkryć bezspornych, obiektywnych moralnych faktów lub na drodze czystego rozumu podać konkretnych, praktycznych rozwiązań. Nie można przedstawić niezależnego od materialnych opinii dowodu, iż dana ocena jest autentycznie prawdziwa, gdyż koresponduje z realnym porządkiem wartości. Niemożliwe są empiryczne dowody prawdziwości ocen. Pozostaje jedynie próba koordynacji ludzkich interesów, przy założeniu, że spełni ona formalne warunki narzucone przez język. Jednak aby droga taka była możliwa, w praktyce musimy prowadzić moralny dialog, dowiadywać się, co dla innych wydaje się pożądane i dobre. Deskrytywizm natomiast, definiując pojęcia moralne poprzez zmienne w czasie i przestrzeni kryteria ich aplikacji, uniemożliwia osiągnięcie tego celu. Zdaniem Hare'a, jedynie za pomocą myślenia krytycznego możemy rozsądzić, jaki kształt mają przybrać zasady moralne, które nazwiemy prawdziwymi. To myślenie krytyczne powinno umożliwić nam udzielenie odpowiedzi na pytanie, czy standardy i obowiązujące warunki prawdziwości są tymi, które powinniśmy zaakceptować. Z tego względu etyka normatywna nie może jako przesłanek używać standardów oceniania. Adekwatna teoria metaetyczna powinna zatem przedstawić neutralną wobec zastanych opinii moralnych koncepcję znaczenia, niezależną od zmiennych i różnych kryteriów aplikacji ocen, oraz umożliwić w jej ramach spójne funkcjonowanie skutecznej etyki normatywnej (§ 3.4). Zdaniem Hare'a, żadna wersja deskrytywizmu nie jest zdolna do spełnienia tego warunku.

¹ Zob. P. Railton, *Moral realism*, „Philosophical Review” (95) 1986, s. 163-207.

² Zob. T. Nagel, *The Possibility of Altruism*, Clarendon Press, Oxford 1970.

³ Zob. J. McDowell, *Values and Secondary Qualities*, w: T. Honderich (ed.), *Morality and Objectivity*, Routledge & Kegan Paul, London 1985.

⁴ Zob. M. Smith, *The Moral...*, dz. cyt.

⁵ Zob. P. Railton, *Moral...*, dz. cyt.

⁶ Zob. M. Smith, *The Moral...*, dz. cyt., s. 11-12.

⁷ Zob. H. Putnam, *Reason, Truth, and History*, Cambridge University Press, New York 1981.

Przeglądając jednak współczesne stanowiska deskrytywistyczne i możliwości ich rozwoju, stwierdzić należy, iż teza, że deskrytywizm prowadzi do relatywizmu, choć w ostatnich latach broniona również przez M. Timmonsa i T. Horgana, jest zbyt mocna. Jak bowiem wskazaliśmy w rozdziale (§ 3.3.2), na podstawie pewnych przekonania dotyczących filozofii języka Hare założył, że znaczenie pojęć moralnych w teoriach deskrytywistycznych musi być zależne od zmiennych i niepowszechnych kryteriów stosowania pojęć moralnych, co uniemożliwia wyjaśnienie istnienia autentycznych sporów moralnych. Istnieją jednak w ramach deskrytywizmu dobrze uzasadnione sposoby uniknięcia również tego, w mniemaniu Hare'a najcięższego, argumentu przeciwko deskrytywizmowi.

Ocena teorii normatywnej Hare'a napotkać musi jeszcze większe przeszkody, niż jest to udziałem jego metaetyki. Wynikają one przede wszystkim z braku szeroko akceptowanych kryteriów dla takiej ewaluacji. Trudno jest wskazać bowiem w miarę niezależną od konkretnych uwarunkowań moralnych, światopoglądowych, ideologicznych czy religijnych metodologię etyki: teorię, która z archimedesowego punktu oparcia pozwalałaby rozsądzić między alternatywnymi koncepcjami normatywnymi.

Przy ocenie utilitaryzmu Hare'a ekwiwalentem kryterium doświadczenia, którego użyliśmy jako jednego z narzędzi oceny teorii metaetycznej, mógłby być zbiór truizmów czy materialnych „prawd moralnych”. Ich status w tej roli wydaje się być jednak wątpliwy. Każda dobra teoria etyczna powinna zabraniać tak oczywistych zachowań, jak np. zadawania bólu osobie bezbronnej i niewinnej dla zabawy, bez widoku na społeczne korzyści, jednak tego typu truizmy nie stanowią problemu dla większości konkurentek utilitaryzmu Hare'a. Bardziej wymagające intuicje moralne są natomiast często przedmiotem kontrowersji i same powinny być uprawomocnione lub skrytykowane przez rzetelną teorię etyczną. Wszak to teoria normatywna ma stanowić probierz wyboru między skonfliktowanymi moralnymi przekonaniem. Z tego względu przy poniższej próbie oceny nie będziemy brali pod uwagę konkretnych materialnych sądów i zasad moralnych, których teoria broni lub które są z nią sprzeczne.

Jedną z najważniejszych zalet etyki normatywnej Hare'a stanowi fakt, że pomimo pewnych trudności wydaje się być ona praktyczna zarówno w wymiarze pedagogicznym, osobistym, jak i społecznym,

co jej autor udowodniał szerokim zakresem podejmowanych problemów praktycznych. Na niewątpliwe wyróżnienie zasługuje również to, iż starał się on uzasadnić swój utilitaryzm na podstawie tezy metaetyki (§ 6.2). Powoduje to, że jego teoria jest spójną całością, w której oba elementy (metaetyka i etyka normatywna) uzyskują ugruntowanie w sobie nawzajem, nie będąc jednak powiązane ze sobą w sposób konieczny. Jeśli założymy bowiem instrumentalną teorię racjonalności, w systemie Hare'a uniwersalność prowadzi do bezstronności, a preskryptywność do zważania na preferencje. Dokładając przekonanie, że najlepszą metodą osiągnięcia wiedzy o preferencjach innych ludzi, zważenia ich siły i próby maksymalizacji ich spełnienia jest wyobrazeniowe postawienie się na ich miejscu, otrzymujemy Hare'a metodę myślenia krytycznego. Dalsza analiza kondycji ludzkiej, naszych wad i słabości motywuje natomiast podział etyki na dwie płaszczyzny: krytyczną i intuicyjną.

Ponadto jego etyka wydaje się być zgodna z naturalistycznym obrazem świata czy faktami naukowymi: nie postuluje istnienia władz poznawczych ani bytów, które trudno jest pojąć na gruncie współczesnej wiedzy pozafilozoficznej. Kolejną jej niewątpliwą zaletą jest pewna doza racjonalizmu: namysł moralny jest w znacznej mierze funkcją argumentów, intersubiektywnie sprawdzalnych rozumowań (§ 6.3.4). Cecha ta pozwala przezwyciężyć relatywizm, który wydaje się być wadą niektórych koncepcji konkurencyjnych. Na uwagę zasługiwać może również to, iż utilitaryzm Hare'a jest przeformułowaniem i rozwinięciem długiej tradycji etyki opartej na tzw. złotej regule (§ 5.2). W znacznym stopniu szanuje również zastane zasady moralne, intuicje oraz wagę pojęć cnoty i wady, uzasadniając ich występowanie w codziennym namyśle moralnym (§ 7). Podaje wyjaśnienia, dlaczego takie cnoty, jak wrażliwość, wyobraźnia, umiejętność przewidywania, bezstronność – cechy zakładane przez tak ważne dla wychowania moralnego pytania, jak np.: „Jak byś się czuł, gdyby tobie ktoś zrobił krzywdę?” – istotne są w edukacji i codziennej praktyce moralnej (§ 7.4). Warto również wymienić jej zdolność do przezwyciężania lub neutralizowania najważniejszych filozoficznych zarzutów skierowanych w stronę klasycznych teorii utilitarystycznych jak i bezpośrednio pod jej adresem (§ 7.1, § 7.2, § 7.3).

Owa skuteczność w odpieraniu zarzutów może nie stanowić jednak tak dużej wartości, gdy zauważymy, iż wiele elementów etyki

Hare'a wymaga doprecyzowania i uszczegółowienia, wyrażona jest bowiem w dość ogólnej i metaforycznej formie. Jego utylitaryzm swą moc przewyższania zarzutów może zawdzięczać pewnej metaforyczności oraz jedynie sygnalizowaniu rozwiązań. Zakłada on, niejednokrotnie bez podania precyzyjnych kryteriów wyboru czynu słusznego, nierozpisaną na procedury praktyczną mądrość. Przejawia się to w braku rozwiniętej, szczegółowej koncepcji: racjonalnego namysłu; tego, kiedy myśleć jak archanioł, a kiedy jak robot; wyboru zakresu konsekwencji, które powinniśmy brać pod uwagę; zakresu hipotetycznych sytuacji, które są istotne; kryterium relewantnych konsekwencji. Takie braki i niejednoznaczności, na tle naturalistycznych prób unaukowania etyki czynionych przez niektórych współczesnych etyków, mogą wydać się wadą.

Kolejną istotną słabością etyki Hare'a jest jego kapitulacja wobec amoralizmu: zgoda, że można być racjonalnym złoczyńcą (§ 7.14). Fakt ten jest tym bardziej znaczący, iż wielu czytelników Hare'a pokładało w nim nadzieję na odnowienie Kantowskiego projektu ufundowania etyki, jeśli nie na koniecznych zasadach rozumu, to na analitycznych regułach języka. Postulowana przez Hare'a nadrzędność języka moralnego nie uzyskała jednak oczekiwanego charakteru i uzasadnienia. Ostatecznie zasady moralne w świetle jego systemu mają charakter jedynie imperatywów hipotetycznych. Każda konkurencyjna teoria, która uniknie takiej „alienacji” moralności od ludzkiego rozumu, wydaje się być bardziej wartościowa.

Również samej próby ufundowania etyki normatywnej na tezach metaetyki, co powszechnie mu się przypisuje, nie można uważać za sukces. Choć czynione przez niego założenia, które wykraczają poza zasady językowe, nie są nadmiernie kontrowersyjne, nie można powiedzieć, aby nie miały one krytyków. Presuponowanie instrumentalnej racjonalności, tego, że racjonalne działanie maksymalizuje spełnianie preferencji, zakładanie zasady pełnego utożsamienia się z innym, teorii relewancji, nondeskryptywistycznej koncepcji tożsamości osobowej, podważa bardzo ambitny plan, który przejawia się w pracy *The Language of Morals*.

Z powodu sporności cech, które mogłyby stanowić kryterium wyboru etyki normatywnej, a także dokonanych przez Hare'a założeń, niektóre z wyżej ocenionych zalet i wad w oczach niejednej osoby mogą stanowić swoje przeciwieństwo. Warto wskazać jednak

również te z presupozycji teorii Hare'a, które choć ważne, jeszcze trudniej jest jednoznacznie osądzić. Taki charakter ma pewien zbiór liberalno-chrześcijańskich wartości, które są podzielane przez znaczą część społeczeństw europejskich i które są widoczne w jego teorii. Wśród nich znajdują się: welferystyczne przekonanie, że nadrzędną wobec grupy jest jednostka z jej preferencjami; liberalna presupozycja, iż nie istnieje wspólny wszystkim ludziom model szczęśliwego życia, zatem nie można budować etyki na powszechnych potrzebach, lecz jedynie na różnorodnych preferencjach; oświeceniowa wiara w skuteczność i prawomocność autorytetu ludzkiego rozumu w stanowieniu zasad moralnych; chrześcijańska wiara, że wszyscy jesteśmy równi, która realizuje się poprzez Benthamowską regułę: „każdy ma liczyć się jako jeden i tylko jako jeden”; oświeceniowe założenie, że jeśli znajdziemy się w idealnej sytuacji epistemicznej, wszystkie problemy moralne możemy skutecznie rozwiązać za pomocą rozumu. Każda z powyższych tez może stanowić źródło krytyki, w oczach innych natomiast być wartością, ze względu na którą warto bronić i rozwijać teorię moralną Richarda Marvina Hare'a.