

Science (of) Fiction

Zur Zukunft des Gedankenexperiments in der Philosophie des Geistes¹

Daniel Cohnitz
Department of Philosophy
University of Tartu

Egal was der heutige Tag auch bringen mag, der 1. April 2063 wird zumindest als der Tag in die Geschichte des Wissenschaftsjournalismus eingehen, der die bisher aufwändigste Berichterstattung erfahren hat. So viele Kamerateams, wie hier vor den Toren der *Australian National University* in Canberra, hat bisher kein wissenschaftliches Experiment anziehen können. Selbst der Knüller des Vorjahres, als es einer 48jährigen Hausfrau in einem Vorort von London gelang, mit einfachsten Küchenutensilien einen kleinen Kalte-Fusion-Reaktor herzustellen, der den Staubsauger und die Mikrowelle zuverlässig mit sauberer Energie versorgte, hatte bei weitem nicht ein solches Echo in der Wissenschaftspresse erfahren.

Es war zwar noch nichts über den Experimentausgang zu erfahren, trotzdem wurde freilich seit drei Tagen ununterbrochen live auf fast allen Kanälen aus Canberra darüber berichtet, dass nun zwar noch nichts zu erfahren sei, dass es aber ja sehr bald soweit sein müsse. Das Experiment, dessen Ausgang man hier so gespannt erwartete, war schon im Vorfeld hitzig diskutiert worden und überhaupt nur durch einen Sonderbeschluss der Vereinten Nationen zustande gekommen. Immerhin läuft die Vorbereitung bereits seit fast 30 Jahren, als sich mehrere Philosophen und Bürgerinitiativen zusammenschlossen, die UNO um eine endgültige Klärung einer der ältesten Fragen der Menschheit zu bitten: Was ist Bewusstsein? Und insbesondere: Ist Bewusstsein etwas anderes als physikalisch beschreibbare Zustände unseres Gehirns, oder sind alle Tatsachen letztlich physikalische Tatsachen? Dass diese Fragen trotz ihres Ehrfurcht einflößenden Alters zunehmend auf den Nägeln brannten, hatte nicht zuletzt die jüngste Entwicklung der Computertechnologie befördert. Nachdem die KI in den letzten Jahrzehnten einen großen Schritt vorwärts gemacht hatte, war man in Sorge, dass ohne endgültige begriffliche Klärung dieser Fragen unser zukünftiger Umgang mit unseren elektronischen Helfern auf moralisch unsicherem Boden stehen könne.

Man hatte sich daher nach langer Abwägung der ethischen Aspekte dieses Experiments dazu entschieden, ein gerade neugeborenes Waisenkind in einem völlig schwarz-weiß-grauen Raum groß werden zu lassen, sorgsam darum bemüht, dass die kleine Mary (so hatte man das Mädchen in Anlehnung an die Protagonistin einer fiktiven philosophischen Geschichte genannt) keine Farben zu Gesicht bekommt. Zugleich ließ man Mary durch schwarz-weiß maskierte Betreuer und schwarz-weiße Lehrvideos eine vorbildliche Ausbildung zukommen. Mit ihren nun 24 Jahren ist Mary bereits in Wahrnehmungspsychologie promoviert und eine absolute Expertin für alle neurophysiologischen Aspekte der Farbwahrnehmung, ein Forschungsgebiet, das in den letzten 10 Jahren stagniert, seit man dort der Meinung ist, alles, was es über diesen Bereich zu wissen gibt, bereits in Erfahrung gebracht zu haben. Heute ist nun der große Tag, an dem Mary zum ersten Mal eine rote Tomate zu Gesicht bekommen soll. Von Marys Reaktion auf diesen Wahrnehmungsreiz erhofft man sich tiefe Einblicke in die Natur des phänomenalen Bewusstseins: Wird Mary überrascht sein, wenn sie zum ersten Mal etwas Rotes sieht? Wird sie dabei etwas Neues über Wahrnehmungspsychologie lernen,

¹ Ich möchte Stefan Bagusche und Myung Hee Theuer für scharfsinnige Kommentare zu früheren Versionen des Textes herzlich danken. Einige Abschnitte dieses Textes sind eng an Cohnitz 2006 angelehnt; dort wird dieses Thema ausführlicher behandelt. Die vorliegende Arbeit wurde unterstützt durch ETF 7163 (*Bridging Explanatory Gaps*) der *Estonian Science Foundation*.

das sie aus ihren schwarz-weißen Lehrvideos alleine nicht hatte lernen können? Neodualisten und aufrechte Materialisten sind gleichermaßen gespannt auf den Experimentausgang; immerhin wird eine der beiden Gruppen ab morgen besser die Profession wechseln, wird doch eine These, die man über Jahrhunderte (wenn nicht Jahrtausende) mit der Inbrunst metaphysischer Überzeugung vorgetragen hat, als von jeher und ein für alle mal falsch entlarvt sein. Trotz der Bedrohung, gleich einer enormen Peinlichkeit ausgesetzt zu sein, haben sich alle wissenschaftlichen Größen aus der philosophischen Debatte hier vor der ANU versammelt.

Endlich beginnt die Übertragung aus dem Inneren des psychologischen Labors. Auf Großleinwänden sieht man in das schwarz-weiße Gefängnis der begabten Wissenschaftlerin. Mary hat am ganzen Körper Elektroden angebracht, die ihre neuronalen Aktivitäten, den Hautwiderstand und all das aufzeichnen, was man hoffentlich als Anhaltspunkt dafür nehmen kann, ob der Anblick der roten Tomate eine Überraschung darstellen wird. (Aus Holland importierte, gentechnisch nur unwesentlich manipulierte graue Tomaten hat Mary schon zu Genüge kennen gelernt. Eine etwaige Überraschung wird man also mutmaßlich auf die Farbe zurückführen können.)

Jetzt ist es endlich soweit. Langsam öffnet sich der Schacht, durch den Mary für gewöhnlich ihr farblich verändertes Essen bekommt. Diesmal hält er aber etwas anderes bereit. Mary geht neugierig auf den offenen Schacht zu, dessen Inneres sie noch nicht sehen kann. Da: Sie stutzt. Nun nimmt sie die Tomate aus dem Schacht und sagt laut und deutlich: „Wow, so that is what it is like to see red.“ Ihre weiteren Worte gehen im aufbrausenden Tumult vor dem psychologischen Labor unter. Hier vor der ANU spielen sich nun unglaubliche Szenen ab. Während eine verschwindend kleine Minderheit aus der Gruppe der Materialisten schnurstracks das Gelände verlassen hat, liegen sich die übrigen Materialisten, wie auch sämtliche Neodualisten freudetaumelnd in den Armen. Alle gratulieren sich durcheinander. Wie kann das sein? Also diese Philosophen ... Wie Politiker, die nach verlorener Wahl, das miserable Abschneiden ihrer Partei noch als moralischen Sieg erklären können, scheint nun auch hier doch mindestens eine der beiden Parteien die Konsequenzen des soeben Erlebten nicht akzeptieren zu wollen. Dieses Experiment war doch jahrzehntelanger Zankapfel zwischen Materialisten und Neodualisten; wie können nun *beide* diesen Ausgang als Bestätigung ihrer Position werten? Dabei war die Hoffnung doch so groß, dass man durch eine empirische Entscheidung das ewige Diskutieren über hypothetische, ja an den Haaren herbeigezogene Fälle nun vergessen kann. Doch nichts da: Philosophen sind offenbar unverbesserlich. Verlassen wir diesen Schauplatz.

1. Einleitung

Würde man das Gedankenexperiment von Frank Jackson, auf dem die obige kleine Geschichte mit Mary, der Wahrnehmungsphysiologin, die ohne Farben aufwächst, beruht, auf diese Weise in ein tatsächliches Experiment überführen, wäre das Ergebnis sehr wahrscheinlich so, wie hier beschrieben: Mary würde beim Anblick der roten Tomate etwas lernen (was sie selbst vermutlich so beschreiben würde: dass es so ist, wie es ist, rot zu sehen). Diese Tatsache würde den Streit um das Gedankenexperiment aber keineswegs beenden. Was für ein peinlicher Befund. Bedeutet das nicht, dass Philosophen es einfach mit der Wahrheitsliebe nicht so genau nehmen, und wie der Astrologe, der verbohrt Psychotherapeut und der unwissenschaftliche Marxist alle widerstreitende Evidenz immer in eine Bestätigung ihrer Theorie umdeuten?

In einer vergleichenden Studie² zu Gedankenexperimenten in den Naturwissenschaften und der Philosophie, in der zunächst festgestellt wurde, dass der Streit um *naturwissenschaftliche* Gedankenexperimente zumindest manchmal durch eine reale Umsetzung des Experiments

² Eine umfassende Kritik dieser Studie findet sich in Cohnitz 2006a.

entschieden werden könne, stellen Jeanne Peijnenburg und David Atkinson mit Blick auf die Gedankenexperimente in der Philosophie folgendes fest:

[I]t is unclear how we ever could put the Chinese Room or the Mary experiment to the test. To be sure, both the Chinese Room and the Mary experiment can be carried out, ethical considerations aside, but that would not resolve the philosophical conundrum. (Peijnenburg und Atkinson 2003, 317)

Und in einer Fußnote hierzu führen sie aus:

This has been shown by Marjolein Degenaar for another thought experiment, viz. Molyneux's problem, named after the Irish philosopher William Molyneux, who presented it in 1688 to John Locke (Degenaar 1996). Imagine a man, blind from birth and capable of distinguishing and naming a globe and a cube by touch. Suppose that this man suddenly acquires sight. Would he be able to distinguish and name both objects simply by looking at them? The question provoked passionate discussions between 18th century rationalists and empiricists, who answered it with yes and no respectively. As Degenaar notes in her excellent book on the subject, the discussion took a new turn after the first cataract operations. From a pure thought experiment, Molyneux's problem turned into a question that could be answered on the basis of real experimentations. Surprisingly enough, however, no agreement was reached. Rationalistic and empiristic scholars kept harassing each other, now over the correct interpretation of the experiment. Degenaar concludes that Molyneux' problem cannot be solved empirically. If she is right, then the Chinese Room and the Mary experiment can certainly not be resolved. (Peijnenburg und Atkinson 2003, 321, FN 7)

Was für ein Skandal! Sind Philosophen also seit jeher nur unkritische Ignoranten? Die wissenschaftstheoretische Literatur, die sich mit dem Thema Gedankenexperiment in der Philosophie beschäftigt, kommt häufig zu genau diesem Ergebnis: Gedankenexperimente in der Philosophie seien fast immer bloßer Ausdruck der eigenen theoretischen Überzeugung. Die „Intuition“, die im Gedankenexperiment als „Gegenbeispiel“ gegen die gegnerische These herangezogen wird, ist schlicht die Intuition, dass die gegnerische These eben falsch ist – eine Intuition, die freilich vom Gegner nicht geteilt wird, weshalb Gedankenexperimente so selten überzeugend sind. Dass eine Ausführung des Gedankenexperiments keinen Fortschritt bringt, liegt daran, dass es hier offenbar gar nicht um Wahrheitsfindung, sondern um die Verteidigung fundamentaler Weltanschauungsdifferenzen geht. Will Philosophie also jemals einen Schritt vorwärts machen, sollte sie sich von der Methode des Gedankenexperiments zukünftig verabschieden. Debatten, in denen Gedankenexperimente als Argumente ausgetauscht werden, müssen dauerhaft auf der Stelle treten, da sich die Debattierenden mit ihren „Argumenten“ eben nur wechselseitig versichern, dass sie intuitiv der Meinung sind, sie selbst hätten recht.

Doch dieses Bild ist bestenfalls eine Verzerrung der Tatsachen. Natürlich gibt es in der Philosophie, wie in anderen Wissenschaften auch, Diskussionen, die sich für eine gewisse Zeit im Kreise drehen, bis die Beteiligten sich dessen bewusst werden und andere Strategien entwickeln. Auch gibt es in der philosophischen Fachliteratur ohne Zweifel Gedankenexperimente, die in dem oben beschriebenen Sinne nur dazu geeignet sind, der Intuition, dass die zur Diskussion stehende These falsch ist, Ausdruck zu geben, ohne aber einen weiteren Grund für ihre Zurückweisung zu präsentieren. Vertreter der fraglichen These teilen diese Intuition dann entsprechend auch nicht.

In den weitaus meisten Fällen funktionieren Gedankenexperimente in der Philosophie aber anders. Kehren wir zu Mary zurück: Mary ist eine junge Frau, die in einem schwarz-weißen

Raum aufwuchs und nie eine (andere) Farbe zu Gesicht bekam. In ihrer Zeit im Gefängnis hat sie alles gelernt, was sich aus neurowissenschaftlicher Sicht zum Farbsehen sagen lässt (da angenommen wird, dass Mary in einer Zukunft lebt, in der die Neurowissenschaft in diesem Sinne abgeschlossen ist). Mary weiß, was in ihren Mitmenschen passiert, wenn diese eine Farbe wahrnehmen, insbesondere weiß sie auch, in welchem Gehirnzustand man sich befindet, wenn man die Farbe Rot wahrnimmt. Nun darf sie zum ersten Mal ihr Gefängnis verlassen und ihr Blick fällt auf einen typischerweise roten Gegenstand. Mary bildet nun einen Gedanken, den sie vorher noch nie hatte: ‚So ist es also, wie es ist, rot zu sehen.‘

Das intuitive Urteil, das hierbei provoziert werden soll, ist das Urteil, dass es sich bei diesem Gedanken von Mary um neues Wissen handelt – sie hat etwas gelernt, was sie aus den schwarz-weißen Lehrbüchern und Videos nicht hätte lernen können. Obwohl eine Expertin für Farbwahrnehmung, gab es dennoch etwas über Farbwahrnehmung, was ihrer Kenntnis bisher entgangen sein musste und was durch ihren neuen Gedanken ausgedrückt wird. Weiterhin handelt es sich um Wissen von einer Tatsache, die auch schon vorher bestanden hat, die also nicht erst durch Marys Wahrnehmungsakt entstanden ist. Wie aber kann dann die Auffassung des Physikalisten wahr sein, dass alle Tatsachen bezüglich Wahrnehmung letztlich physikalische Tatsachen sind, wenn es Tatsachen bezüglich Wahrnehmung gibt, die jemandem, der perfekt über alle physikalischen Tatsachen informiert ist, entgangen sein können? Der Physikalismus muss also falsch sein.

Letzteres ist eine Folgerung aus dem Ergebnis des Gedankenexperiments und Annahmen darüber, was für Phänomene eigentlich in den Erklärungsbereich des Physikalismus fallen. Die intuitive Beurteilung des Gedankenexperiments beschränkt sich auf die Frage, ob man Marys Überzeugung ‚So ist es also, wie es ist, rot zu sehen.‘ als ‚neues Wissen‘ beschreiben soll oder nicht. Die Intuition, die hier zur Debatte steht, bezieht sich nicht darauf, ob der Physikalismus falsch ist. Dass er falsch sei, folgt überhaupt erst, wenn man zusätzliche weitere Annahmen macht.

In der Tat gibt es Philosophen, die angezweifelt haben, dass das Gedankenexperiment unproblematisch etablieren kann, dass Mary hier neues Wissen erwirbt. Schließlich sei unklar, was es heißen soll, dass die Neurowissenschaft, in der Mary ausgebildet wird, ‚vollständig‘ ist. Entweder, ‚vollständig‘ heißt soviel wie ‚alles, was man über Farbwahrnehmung wissen kann‘, dann kann Mary ja *per definitionem* nichts Neues gelernt haben. Oder es heißt soviel wie ‚alles, was man über Farbwahrnehmung wissen kann, bis auf das, was Mary nun neu hinzulernt‘ (oder: ‚... bis auf den qualitativen Aspekt phänomenaler Erfahrungen‘), dann würde in die Beschreibung des Gedankenexperiments aber eine Annahme eingehen, die erst zu zeigen wäre, es handelte sich also um eine *petitio*. Hierbei werden keine anders gelagerten Intuitionen herbeizitiert, sondern die Kritik bezieht sich darauf, ob das Gedankenexperiment durch seine Beschreibung nicht an Aussagekraft verliert.

Andere Philosophen haben dieses Gedankenexperiment kritisiert, indem sie darauf hinwiesen, dass der Übergang von ‚Mary hat neues Wissen erworben‘ auf ‚Es gibt eine weitere Tatsache über Farbwahrnehmung, die in der neurowissenschaftlichen Beschreibung nicht enthalten war.‘ unzulässig ist. Diese Philosophen analysieren Marys neues Wissen als ‚*knowing how*‘, anstatt als ‚*knowing that*‘, also als eine Form nicht-propositionalen Wissens, dem keine eigene Tatsache entsprechen muss. Oder sie analysieren Marys neues Wissen als Wissen von einem neuen Aspekt bezüglich einer Tatsache, die ihr unter anderen Hinsichten bereits bekannt war. Auch diese Strategien warten nicht mit anderen Intuitionen auf. Die Intuition, dass Mary etwas Neues gelernt hat, wird auch von diesen Philosophen geteilt, was bezweifelt wird, ist die Frage, ob es sich bei diesem neuen Wissen um Wissen von einer neuen Tatsache handelt, es geht also darum, welche Folgerungen man aus der Intuition ziehen darf. Diese Debatte wird geführt, indem man verschiedene Theorien über *knowing how* und *knowing that* miteinander vergleicht, bzw. verschiedene Theorien über propositionale Einstellungen heranzieht, diese im Lichte unabhängiger Daten miteinander vergleicht, die plausiblere auf

den Fall appliziert und zusieht, was dabei herauskommt. Natürlich benutzen Philosophen auch dabei wieder Gedankenexperimente und befragen wiederum ihre Intuitionen, allerdings sind es diesmal keine Intuitionen über Mary und ihr Wissen. Die Frage, ob es sich bei Marys neuem Wissen um *knowing how* oder *knowing that* handelt, ist von der Physikalismus-Frage unabhängig.

Aber was ist mit den Philosophen, die tatsächlich dem Gedankenexperiment vorwerfen, es sei zirkulär? Haben wir nicht einige Sätze zuvor eingeräumt, dass manche Philosophen meinen, dass das Gedankenexperiment in seiner Beschreibung voraussetzt, was eigentlich zu zeigen ist? Ja, dem ist so, aber das scheint in keiner Weise für die Methode des Gedankenexperiments *an sich* problematisch zu sein. Alle an der philosophischen Diskussion Beteiligten sind sich einig darüber, dass das Gedankenexperiment schlecht wäre, wenn es in seiner Beschreibung voraussetzen würde, was eigentlich zu zeigen ist. Deshalb findet man solche Vorwürfe auch *innerhalb* der philosophischen Debatte. Wenn es in der Philosophie zum normalen Geschäft gehörte, dass Gedankenexperimente eine *petitio principii* begehen, gäbe es kein solches Kontrollinstrument innerhalb der Philosophie des Geistes. Gedankenexperimente wären doch nur dann ein problematisches Instrument in der Philosophie, wenn es ausschließlich den Wissenschaftstheoretikern auffiele, wenn sie eine *petitio* begingen, alle (oder die Meisten) an der Debatte Beteiligten dies aber für normal und untadelig hielten.

Gedankenexperimente sind – auch in der Philosophie – im Allgemeinen nicht zirkulär. Es wird von den an der Debatte Beteiligten genau darauf geachtet, ob ein Gedankenexperiment in seiner Beschreibung das zu Beweisende bereits vorwegnimmt, ob die fragliche Intuition tatsächlich allgemein geteilt wird und ob die Tatsache, dass das Gedankenexperiment diese Intuition hervorruft, auch irgendwelche Aussagekraft für die Konklusion hat, die in dem Gedankenexperiment-Argument gezogen wird.

Diese Debatten werden (zum Teil) geführt, ohne dass empirische Versuche angestellt würden. Philosophische Gedankenexperimente werden normalerweise nicht empirisch realisiert. Auch das hatte unsere kleine Eingangsgeschichte bereits aufgezeigt: Eine empirische Untersuchung wird nicht nur selten angestrebt, eine empirische Untersuchung der Umstände im Gedankenexperiment würde die Debatte vielmehr qualitativ in keiner Weise beeinflussen. Aus den soeben angestellten Überlegungen ist schon halbwegs ersichtlich, warum dem so sein sollte: Wenn der Ausgang des Gedankenexperiments (Mary lernt etwas Neues) gar nicht der Streitgegenstand ist, sondern vielmehr die Frage zu klären ist, wie man diesen Befund zu interpretieren hat, wird natürlich auch eine empirische Etablierung desselben Punkts keinen entscheidenden Fortschritt bringen können.

Trotzdem ist damit natürlich noch nicht der Sinn solcher Gedankenspielerien verteidigt. Selbst wenn es Philosophen gelingt, ihren Austausch abgedrehter Szenarien auf eine zumindest nicht offensichtlich zirkuläre Weise zu betreiben, ist damit nicht gesagt, *warum* sie dies tun, und die offenbare Unabhängigkeit dieses Verfahrens von der Empirie ist weniger Vertrauen erweckend als vielmehr Besorgnis erregend:

So lehrreich, unterhaltsam und manchmal auch vergnüglich derlei Gedankenspiele sein mögen – so irritierend sind sie für diejenigen, die nichts weiter wollen als die wirkliche Welt verstehen, in der wir zu leben glauben. Natürlich rechtfertigen Philosophen derartiges Philosophieren damit, dass auf diese Weise auch die Begrifflichkeit, die wir zur Erfassung der wirklichen Welt haben, geschärft wird. Der empirischen Seite will das allerdings nur schwer einleuchten. Denn warum sollten Erfindungen möglicher Welten und Wesen unsere Begriffe für die wirkliche Welt und wirkliche Wesen schärfen, wenn diese Erfindungen in beliebige Richtungen gehen können? Was sie gewiss schärfen können, sind Sprachregelungen und Begrifflichkeiten zur

Charakterisierung möglicher Welten. Aber was soll die Empirie damit nur anfangen?
(Prinz, in diesem Band, ???)

Doch selbst hierzu gibt es eine gute Antwort. Um genau zu sein, gibt es sogar mehrere gute Antworten. Bei den Antworten, die uns im Folgenden besonders interessieren, werden wir davon ausgehen, dass die Philosophie zumindest prinzipiell an der Erforschung desselben Gegenstandsbereichs interessiert ist wie beispielsweise die Psychologie. Eine solche Annahme muss man natürlich nicht machen und kann demnach die Frage nach dem Zweck philosophischer Gedankenexperimente für den Empiriker einfach zurückweisen: Es gibt keinen solchen Zweck, außer der Empiriker interessiert sich eben auch für philosophische Fragen.³

Wir gehen also davon aus, dass die Philosophie des Geistes ein Unternehmen ist, das von seinem Selbstverständnis her gemeinsam mit der Psychologie und den Neurowissenschaften (und gegebenenfalls weiteren Disziplinen wie der Informatik) den (menschlichen) Geist wissenschaftlich verstehen will. Das beinhaltet zumindest einen moderaten Naturalismus: Ehemals rein philosophisch verstandene Fragen können von den Naturwissenschaften beantwortet werden. Wenn man annimmt, dass dem so ist, welche Rolle spielen dabei noch Gedankenexperimente? Empirische Fragen untersucht man doch mit empirischen Verfahren und nicht durch reines Nachdenken. Außerdem beschäftigt sich die Naturwissenschaft mit den empirischen Zusammenhängen in *dieser* Welt – was soll die Beschäftigung mit möglichen Wesen und Welten?⁴

2.1 DAS GEDANKENEXPERIMENT ALS EINSICHT IN NATURZUSAMMENHÄNGE

Die erste Antwort auf die letzte Frage beginnt mit der Beobachtung, dass naturwissenschaftliche Aussagen, insbesondere Aussagen über gesetzesartige Zusammenhänge, nicht nur Aussagen über das Hier und Jetzt sind. Eine naturwissenschaftliche Theorie redet nicht nur über tatsächlich gemachte Beobachtungen oder tatsächliche zukünftige Beobachtungen, sondern auch über mögliche Beobachtungen und auch über mögliche Sachverhalte, die prinzipiell unbeobachtbar sind. Das ist es (unter anderem), was „gesetzesartig“ zu sein bedeutet: Die wahre Verallgemeinerung ‚Alle Äpfel in diesem Korb sind grün.‘ ist nicht gesetzesartig, weil sie nichts darüber sagt, was der Fall wäre, würde ich diesen Apfel ebenfalls in diesen Korb legen. Bloß akzidentell wahre Verallgemeinerungen schweigen sich über andere mögliche Welten aus. Die gesetzesartige Aussage ‚Alle Gegenstände, die aus Metall sind, leiten Strom.‘ sagt aber sogar etwas darüber, was der Fall wäre, wenn dieser Apfel aus Metall wäre. Mögliche Welten und Wesen sind also schon deshalb von Relevanz für den Empiriker, weil der Gehalt seiner empirischen Hypothesen eben über die aktuelle Welt hinausgeht.

Die schwierige Frage ist weniger, warum diese möglichen Welten und Wesen von Relevanz für den Empiriker sind, sondern eher, warum man meint, mit *Gedankenexperimenten* zu zuverlässigen Aussagen über mögliche Welten und Wesen kommen zu können. An dieser

³ Es geht also im Folgenden darum, die methodische Rolle des Gedankenexperiments für eine mehr oder weniger naturalistisch verstandene Philosophie (des Geistes) zu untersuchen. Zur Methode des Gedankenexperiments in anderen Konzeptionen vgl. Cohnitz 2006.

⁴ Wir beschränken uns im Folgenden außerdem auf die *kritischen* Funktionen von Gedankenexperimenten. Man redet auch von „Gedankenexperimenten“, wenn Konsequenzen einer Theorie zu didaktischen Zwecken an einem hypothetischen Beispielfall verdeutlicht werden. Gegen Gedankenexperimente dieser Art ist – soweit mir bekannt – noch keine Kritik laut geworden. Ihre Evaluation fällt auch nicht in den Bereich der Wissenschaftstheorie, sondern in den der Didaktik. Zu den verschiedenen Dingen, die man sonst noch so als „Gedankenexperiment“ bezeichnet, vgl. Cohnitz 2006.

Stelle kann man bestenfalls auf die eingeschränkte Zuverlässigkeit unserer empirischen Intuitionen verweisen: In Bezug auf empirische Phänomene, mit denen wir in unserem Alltag häufig zu tun haben (bzw. unsere UrUrUrahnen in ihrem Alltag häufig zu tun hatten), können wir (in gewissem Rahmen) zuverlässig intuitive Auskünfte geben. Dies kann sehr hilfreich sein, solange eine Wissenschaft sich noch in einem Frühstadium befindet, in dem nur wenige empirische Daten über den Gegenstandsbereich gesammelt vorliegen und davon ausgegangen werden kann, dass ein signifikanter Teil des Gegenstandsbereichs dieser Wissenschaft uns tatsächlich bekannt ist. In diesem Fall sind die Ergebnisse von Gedankenexperimenten nicht nur von heuristischem Interesse, sondern rechtfertigen auch die durch sie initiierten Überzeugungen. Es handelt sich dann um *prima facie gerechtfertigte Anfangshypothesen*. Diese Rolle können philosophische Gedankenexperimente auch für die Psychologie spielen:

Perhaps the most obvious application of intuitions to the empirical sciences in general, and those of philosophical intuitions to psychology, in particular, comes when we use our intuitions as a source of empirical hypotheses. We might think of intuitive judgments as particularly plausible initial hypotheses about the nature of the world. [...] Like all such hypotheses, our intuitions are accurate as a guide to the world to different degrees, depending on both the person who has those intuitions and the subject area to which they apply. [...] We are all experts, in a pragmatic sort of way, in everyday psychology, and for this reason our everyday psychological intuitions can be taken as plausible initial hypotheses about how the mind really works. Empirical psychology can then use these everyday intuitions as a starting point for its research. (Gopnik and Schwitzgebel 1998, 78-79)

Die Intuition stattet uns also nicht bloß heuristisch mit Anfangshypothesen aus, wenn „Heuristik“ tatsächlich nur den Entdeckungszusammenhang und nicht den Rechtfertigungszusammenhang einer Hypothese betrifft. Sofern die Intuitionen der richtigen (relevanten) Personen in den richtigen Kontexten zu Hypothesen führen, sind diese Hypothesen aufgrund ihrer Herkunft auch mit rechtfertigender Anfangsplausibilität ausgestattet. Auch Hilary Kornblith – ein prominenter Kritiker der philosophischen Gedankenexperimentiererei – ist bereit, soviel zuzugestehen:

[W]hy do philosophers spend so much time scrutinizing their intuitions, that is, looking inward, if, on my view, what they are really interested in is external phenomena? [I]f I am asked a question about rocks, for example, one way to answer the question is to ask myself what I believe what the answer is. Although I am asked a question about rocks, I answer it by enquiring into what I believe. This is a perfectly reasonable thing to do if I have good reason to think that my current beliefs are accurate, or if I do not have access to a better source of information. By looking inward I answer a question about an external phenomenon. This, to my mind, is what we do when we consult our intuitions. (Kornblith 2002, 14-15)

Unsere Intuitionen sind umso zuverlässiger, desto näher der fragliche Sachverhalt, der im Gedankenexperiment zur Debatte steht, an unseren vergangenen Erfahrungen ist. Zugleich ist aber auch klar, dass diese intuitiven Urteile über den Gegenstandsbereich mit der fortschreitenden Entwicklung der empirischen Wissenschaft von diesem Bereich in den Hintergrund treten müssen.

The intuitions articulated by philosophers, then, in so far they are treated as initial hypotheses about the way the world really is, can guide empirical research, especially in those domains in which human beings have particular expertise. [...] As sciences

mature, however, they typically revise, alter, and sometimes entirely reject, these initial hypotheses. In Galileo's day, it was permissible for physicists in defending one view or another to appeal to what we would now call our "folk physical" intuitions about what would happen in various circumstances. In contemporary physics, such appeals would be ruled out. Similarly, as psychological science has matured, we have become more confident in leaving intuitions behind, when they conflict with well-supported empirical findings [...]. Our initial hypotheses may be, indeed often are, moderated or defeated by later evidence. So although we may justifiably take reflective psychological intuition as a good preliminary guide, we no longer take it as the final authority about the mind. (Gopnik and Schwitzgebel 1998, 79)

Unsere physikalischen Intuitionen – beispielsweise – können uns heute nicht mehr helfen, wenn es darum geht, Theorienkonflikte zwischen rivalisierenden physikalischen Theorien zu entscheiden, was sie aber zu Galileis Zeiten vielleicht noch konnten. Ein ähnliches Schicksal mag unseren Intuitionen in Bezug auf den Gegenstandsbereich der Psychologie beschieden sein. Intuitionen, die sich auf empirische Sachverhalte beziehen, müssen empirischen Untersuchungen standhalten können oder sind eben zurückzuweisen. Ihnen kommt auf jeden Fall keine *besondere* Autorität zu, die sie vor den Ergebnissen des Empirikers auszeichnen würde.

2.2 GEDANKENEXPERIMENTE ALS DATEN FÜR DIE LINGUISTIK UND PSYCHOLOGIE

Anders sieht die Sache aus, wenn es um semantische Intuitionen geht, also um Gedankenexperimente zum Zwecke der Bedeutungsanalyse. Diesen Intuitionen kommt eine gewisse Autorität zu, allerdings muss man hier zwischen verschiedenen Formen der Bedeutungsanalyse unterscheiden. Zunächst kann man Bedeutungsanalysen als sprachwissenschaftliche Untersuchungen der Bedeutung bestimmter Ausdrücke in einer – mehr oder weniger weit gefassten – Sprechergemeinschaft auffassen. Unser Urteil in Bezug auf imaginäre Fälle im Gedankenexperiment wäre dann nicht notwendigerweise eine empirische Hypothese/Prognose über unser tatsächliches Sprachverhalten in kontrafaktischen Umständen, sondern ein simples Datum unseres tatsächlichen Sprachverhaltens: Wir verwenden den Ausdruck X in kontrafaktischen Sätzen eben so. Sinn und Zweck von Gedankenexperimenten dieser Art wäre es, Daten bereitzustellen, die von der (jeweils relevanten) Theorie über die Bedeutung von X in dieser Sprechergemeinschaft zu erklären wären. Diese Daten könnten – wie bei anderen Datenerhebungsverfahren auch – in Bezug auf ihr Zustandekommen kritisiert werden. Soll die Sprechergemeinschaft größer sein als die der westlichen Philosophen und Philosophinnen, sollte man die Datenerhebung unter restringierteren Bedingungen durchführen als in der herkömmlichen Form der Lehnstuhlreflexion. Man müsste kontrollierte empirische Untersuchungen durchführen. Es sollte dabei ausgeschlossen werden, dass die intuitiven Urteile durch theoretische Überzeugungen „verunreinigt“ sind, kultur- und einkommensabhängig sind, etc.⁵ Falls die Ergebnisse eines Gedankenexperiments dann einer Theorie widersprechen, deren Gegenstandsbereich diese Daten enthält, würde die Theorie falsifiziert.

Dies gilt nicht nur für unsere intuitiven Urteile über Grammatikalität und die Bedeutung von Ausdrücken in kontrafaktischen Zusammenhängen, sondern auch in Bezug auf alle möglichen Überzeugungen, die wir sprachlich zum Ausdruck bringen und die von psychologischen Theorien zu erklären sind. So ist unsere intuitive Alltagsphysik zwar irrelevant für die

⁵ Es scheint, dass diese Faktoren das intuitive Urteil in Bezug auf philosophische Gedankenexperimente beeinflussen können. Vgl. Machery et al. 2004.

moderne Physik, aber relevant für die Psychologie, sofern sie sich mit intuitiver Alltagsphysik als psychologisches Phänomen beschäftigt:

When we consider [intuitions] as evidence, intuitions have some of the quality of indefeasibility that evidence always has. Parallel lines may, or may not, meet, and we may, or may not, have first-person access to our mental states, but it is indubitably true that we believe that parallel lines will not meet and that we believe that we have first-person access to our own mental states. A psychological theory has to explain why we have these intuitive beliefs even if, indeed especially if, the beliefs are quite false. The psychologist cannot simply reject the beliefs if she is treating them as evidence about the mind of the person who has them. (Gopnik and Schwitzgebel 1998, 80)

Auch diese Position ist von einem naturalistischen Blickwinkel betrachtet unproblematisch. Welche Adäquatheitsbedingungen für das Gedankenexperiment gelten, hängt davon ab, wie der Gegenstandsbereich der fraglichen Theorie bestimmt ist. Dasselbe gilt für die Frage, ob man kulturrelevante Unterschiede u. Ä. in das Untersuchungsdesign aufnehmen muss, oder ob ein paar ruhige Minuten im eigenen Ohrensessel ausreichen. Hier gelten zunächst einfach die Standardbedingungen für psychologische Experimente.

2.3 GEDANKENEXPERIMENTE ZUR EXPLIKATION THEORETISCHER ZUSAMMENHÄNGE

Doch Philosophen interessieren sich häufig nicht bloß für die Frage, wie ein Begriff *de facto* in der Umgangssprache (bzw. in einer relevanten Sprechergemeinschaft) verwendet wird, sondern machen es sich zur Aufgabe, den Begriff zu „schärfen“. D.h. sie verfolgen ein revisionistisches Programm: Nach der (philosophischen) Analyse soll der Begriff präziser sein, als er es vorher war. Ein Unternehmen, das in der Regel damit begründet wird, dass präzisere Begriffe für das Betreiben von Wissenschaft unerlässlich sind. Eine solche revisionistische Bedeutungsanalyse bezeichnet man spätestens seit Rudolf Carnaps *Logical Foundations of Probability* (Carnap 1967) als „Explikation“.

Wenn wir uns nun fragen, welche Rolle Gedankenexperimente bei der Explikation spielen, müssen wir zwei verschiedene Rollen unterscheiden, von denen man die eine als „interne“ kritische Rolle von Gedankenexperimenten bezeichnen könnte, die andere als „externe“ kritische Rolle. Die Übergänge sind in fast allen Fällen der Gegenwartsphilosophie fließend. Betrachten wir zunächst den Fall der internen Kritik.

Wenn wir einmal von sehr vielen komplizierenden Faktoren absehen, kann man Gedankenexperimente auch dazu verwenden, die interne Kohärenz und Struktur von Theorien zu untersuchen. Indem man imaginäre Fälle entwirft, die in den intendierten Anwendungsbereich der Theorie fallen und die Begrifflichkeiten der Theorie auf die intendierte Weise anwendet, kann man zu nicht-intendierten Ergebnissen kommen. Es kann sein, dass die Vertreter einer fraglichen Theorie bestimmte Sachverhalte anders beschreiben möchten, als die begrifflichen und theoretischen Festlegungen, die sie in ihrer Theorie getroffen haben, dies zulassen. Das eine Extrem bestünde in der Aufdeckung interner Widersprüche in einer Theorie. Russells Antinomie wäre so ein Fall, bei dem an einem bestimmten Beispiel (der Russell-Menge) gezeigt wird, dass bei einer Anwendung der vorgenommenen theoretischen Festlegungen auf diesen Fall ein Widerspruch abgeleitet werden kann. Das andere Extrem bestünde darin, dass man an Beispielen aus dem intendierten Anwendungsbereich zeigt, dass eine konsistente Anwendung der Theorie auf diesen Fall, es nicht erlaubt, den Fall so zu beschreiben, wie man es eigentlich wollte (etwa weil die begrifflichen Festlegung in der Theorie zu grobkörnig sind und die Unterscheidung, die man treffen wollte, gar nicht zulassen). Damit ein Gedankenexperiment in diesem Sinne funktioniert, muss es einen Fall präsentieren, der im intendierten Anwendungsbereich der

Theorie liegt, und dieser Fall muss mit den fraglichen theoretischen Begrifflichkeiten im intendierten Sinne beschrieben werden.

Kritiker der philosophischen Gedankenexperimentiererei sind oft der Meinung, dass mit den soeben aufgeführten Fällen (Gedankenexperimente zur Entwicklung von Anfangshypothesen, Gedankenexperimente zur Bereitstellung von Daten für die Psychologie, Gedankenexperimente zur internen Kritik) alle hauptsächlich relevanten Anwendungen des Gedankenexperiments abgedeckt sind und dass diejenigen Fälle, die uns philosophisch interessanter erscheinen, nur darauf beruhen können, dass man diese Funktionen durcheinander wirft:

Within the philosophy of mind, there has been a tradition of either accidentally confusing different uses of intuition, or deliberately treating them as identical, and consequently drawing inappropriate conclusions. An unspoken assumption of much argumentation in the philosophy of mind has been that to articulate our folk psychological intuitions, our ordinary concepts of belief, truth, meaning, and so forth, is itself sufficient to give a theoretical account of what belief, truth, meaning, and so forth, actually are. We believe that this assumption rests on an inadequate understanding of the nature of intuition and its appropriate applications, and that it results in errors [...]. (Gopnik and Schwitzgebel 1998, 84)

Alison Gopnik und Eric Schwitzgebel vermuten hier, dass Philosophen mit Intuitionen darüber, was sie glauben, was Bedeutung, Wissen, Überzeugung, etc. sind, psychologische Theorien zu kritisieren versuchen. Dies sei aber eine unzulässige Verwendung des Gedankenexperiments. Da die Psychologie aus der Phase herausgetreten ist, in der es noch von Interesse war, zu erforschen, was man denn so über den Gegenstandsbereich für Meinungen hat, sind philosophische Intuitionen heute bestenfalls als empirisches Datum zur Kenntnis zu nehmen, das ggf. von einer Theorie erklärt werden müsste. Dies muss aber nicht notwendigerweise durch dieselbe Theorie geleistet werden, die die Philosophen mit ihren Intuitionen angreifen. Wenn Psychologen beispielsweise der Meinung sind, dass sie mit einem anderen Wissensbegriff oder einem anderen als dem „supraindividuellen“, externalistischen Bedeutungsbegriff besser klar kommen, kann es keinen Grund geben, diese Verwendung aus philosophischer Sicht zu kritisieren:

The fact that the ordinary concept of meaning within our daily experience has [the character of a supraindividual concept] provides only rudimentary evidence about what notion of meaning will work best for psychologists and others who wish to talk about meaning in a scientifically informed way—just as our ordinary intuitions about physics provide only a rudimentary starting point for the creation of physical laws. These intuitions may have played a crucial role for Galileo, but most of them did not survive in the end. (Gopnik and Schwitzgebel 1998, 85)

Haben die beiden nicht Recht? Ist es uns nicht jedesmal peinlich, darauf hingewiesen zu werden, mit welcher Inbrunst sich Kant für bestimmte synthetische Urteile a priori eingesetzt hat, nur um dann durch die Physikgeschichte widerlegt zu werden? Wollen wir das etwa auch riskieren, indem wir das junge Pflänzlein einer interdisziplinären Kognitionswissenschaft dadurch riskieren, dass wir den fleißigen Psychologen mit ausgedachten Experimenten „nachweisen“, sie redeten gar nicht über ‚Wissen‘ und ‚Bedeutung‘? Es hat in wissenschaftlichen Zusammenhängen oft keinen Wert, sich an den Alltagsgebrauch eines Begriffs zu klammern, wenn er in einer anderen Bedeutung fruchtbarer verwendet werden kann.

Wenn Ausdrücke, die in der Alltagssprache bereits vorkommen, in der Wissenschaftssprache mit einer neuen Bedeutung versehen werden, hat dies offensichtlich etwas mit Begriffsexplikationen im Carnapschen Sinne zu tun. Während so genannte *Nominaldefinitionen* neue Ausdrücke mit eigener Bedeutung einführen, und so genannte *Bedeutungsanalysen* die tatsächliche linguistische Bedeutung bereits verwendeter Ausdrücke feststellen, bezeichnet die *Explikation* ein Verfahren, bei dem ein bereits verwendeter Ausdruck (aufgrund wissenschaftlicher Erfordernisse) eine neue, vom bisherigen (außerwissenschaftlichen Alltags-) Gebrauch abweichende Bedeutung erhält. Diese Abweichung des Explicatums (des Ergebnisses einer Explikation) vom Explicandum (dem zu explizierenden umgangssprachlichen Ausdruck) kann natürlich unterschiedlich gravierend ausfallen.

Rudolf Carnap, der das Verfahren der Explikation als erster systematisch beschrieben hat, hatte sich in Bezug auf das Verhältnis von Explikation zu herkömmlicher Begriffsverwendung recht vorsichtig ausgedrückt, da er sich der Tatsache bewusst war, dass theoretische Fruchtbarkeit letztlich viel interessanter ist als die Nähe zur Umgangssprache. Die umgangssprachliche Verwendung – so die Überlegung Carnaps – ist ja gerade ungenau und revisionsbedürftig, sonst bräuchte es ja keine Explikation. Die „neue“ Bedeutung sollte ja gerade abweichen, um diese Ungenauigkeit zu vermeiden.

Wie weit die Bedeutung des Explicatums vom Explicandum abweicht, ist eine Frage der Bedeutungsanalyse und deskriptiv. Sie wird dadurch beantwortet, dass man feststellt, wie ein bestimmter Begriff *de facto* im Alltag verwendet wird. Dass man dabei feststellt, dass die Bedeutung des Explicatums von der Alltagsbedeutung des Ausdrucks abweicht, oder stark abweicht, ist gemessen an den Adäquatheitsbedingungen für Explikationen alleine ziemlich unerheblich. Aus diesem Grunde ist man häufig der Meinung, dass bei Explikationen – ähnlich wie bei Nominaldefinitionen – praktisch alles erlaubt ist, solange man nicht irgendwelche Konsistenzbedingungen verletzt. Es kommt doch letztlich nur darauf an, dass die Wissenschaft schlussendlich über einen „fruchtbaren“ Begriffsapparat verfügt.

Was hierbei übersehen wird, sind die normativen Beschränkungen, die auf die Festlegung des Explicandums wirken. Carnap hatte ganz richtig festgestellt, dass der zu explizierende Begriff unklar, vage und problematisch ist, man also bei der Formulierung der Aufgabenstellung für die Explikation nicht recht sagen kann, was da eigentlich expliziert werden soll:

A problem of explication is characteristically different from ordinary scientific (logical or empirical) problems, where both the datum and the solution are, under favorable conditions, formulated in exact terms. [...] In a problem of explication the datum, viz., the explicandum, is not given in exact terms; if it were, no explication would be necessary. Since the datum is inexact, the problem itself is not stated in exact terms; and yet we are asked to give an exact solution. This is one of the puzzling peculiarities of explication. (Carnap 1967, 4)

Die Aufgabenstellung für eine Explikation kann also nicht genau formuliert werden. Ist es aber nicht ohnehin unerheblich, wie man die Aufgabenstellung genau definiert? Immerhin gibt es doch Adäquatheitsbedingungen (Nähe zur Umgangssprache, Fruchtbarkeit, Einfachheit), die im Nachhinein eine Beurteilung des Explicatums zulassen. Man sollte also meinen, es spiele keine Rolle, wie das Problem der Explikation genau zu formulieren ist.

On the contrary, since even in the best case we cannot reach full exactness, we must, in order to prevent the discussion of the problem from becoming entirely futile, do all we can to make at least practically clear what is meant as the explicandum. What X means by a certain term in contexts of a certain kind is at least practically clear to Y if Y is able to predict correctly X's interpretation for most of the simple, ordinary cases of the

use of the term in those contexts. It seems to me that, in raising problems of analysis or explication, philosophers very frequently violate this requirement. They ask questions like: ‘What is causality?’, ‘What is life?’, ‘What is mind?’, ‘What is justice?’, etc. Then they often immediately start to look for an answer without first examining the tacit assumption that the terms of the question are at least practically clear enough to serve as the basis for an investigation, for an analysis or explication. (Carnap 1967, 4)

Auch bei Begriffsexplikationen ist „alles“ dafür zu tun, dass wir uns wechselseitig darüber „zumindest praktisch verständlich“ machen, wie wir das Explicandum meinen wollen. Ohne eine solche Verständigung wären Explikationsversuche „absolut zwecklos“. Carnap gibt uns auch einen Hinweis darauf, wie eine solche Verständigung erzielt werden kann:

Even though the terms in question are unsystematic, inexact terms, there are means for reaching a relatively good mutual understanding as to their intended meaning. An indication of the meaning with the help of examples for its intended use and other examples for uses not now intended can help the understanding. An informal explanation in general terms may be added. All explanations of this kind serve only to make clear what is meant as the explicandum; they do not yet supply an explication; say a definition of the explicatum; they belong still to the formulation of the problem, not yet to the construction of an answer. [...] By explanations of this kind the reader may obtain step by step a clearer picture of what is intended to be included and what is intended to be excluded; thus he may reach an understanding of the meaning intended which is far from perfect theoretically but may be sufficient for the practical purposes of a discussion of possible explanations. (Carnap 1967, 5)

Zur wechselseitigen Verständigung über das Explicandum werden also Fälle genannt, in denen man das Explicatum benutzen will, und Fälle, auf die es nicht zutreffen soll. Man gibt also Kontexte an, in denen der fragliche Ausdruck in einer wahren Beschreibung vorkommt, sowie Fälle, in denen der Ausdruck nicht Bestandteil einer wahren Beschreibung sein darf. Die Aufgabe der Explikation ist es dann, ein Explicatum zu finden, das die Wahrheitswerte dieser Beschreibungen unverändert lässt, wenn man in ihnen das Explicandum durch das Explicatum ersetzt. In jeder anderen Hinsicht mag das Explicatum abweichen. Die vorher festgelegten Kontexte bestimmen, was man (mindestens) expliziert haben will. Die Anforderung, dass diese Kontexte ihre Wahrheitsbewertungen behalten, wenn man in ihnen das Explicandum durch das Explicatum ersetzt, gehört dann ebenfalls zu den Adäquatheitsbedingungen für die Explikation.

Als Beispiele für solche Kontexte nennt Carnap isolierte Sätze, in denen der Explicandum-Ausdruck vorkommt. Wir sehen aber jetzt, dass in der Philosophie nicht nur isolierte Sätze zur Verständigung über das Explicandum herangezogen werden sondern durchaus reichere Kontexte: Gedankenexperimente sind ebenfalls solche Kontexte. Die Intuitionen, die wir in Bezug auf Gedankenexperimente haben und in der philosophischen Diskussion präsentieren, sind vornehmlich solche, von denen wir der Meinung sind, dass sie als Adäquatheitsbedingungen für Explikationen dienen sollen, und wir gehen implizit davon aus, dass diese Auffassung in der wissenschaftlichen Gemeinschaft geteilt wird. Idealerweise erwarten wir, dass jede zukünftige Begriffsexplikation die intuitive Beurteilung des Gedankenexperiment-Falls erhalten wird. Wir verwenden in diesen Fällen Gedankenexperimente als *Adäquatheitskontexte*:

Definition Adäquatheitskontext: Ein Gedankenexperiment G ist ein Adäquatheitskontext für die Explikation eines Explicandum-Ausdrucks B aus G gdw. jede adäquate Explikation E von B den Wahrheitswert aller Aussagen in G erhält, wenn E in G für B ersetzt wird.

Was dies in der Praxis dann bedeutet, kann man an vielen philosophischen Beispielen sehen. So führt – beispielsweise – Dretske in *Knowledge and the Flow of Information* (Dretske 1999) eine Definition des Wissensbegriffs ein, um dann an klassischen Gedankenexperimenten gegen frühere Explikationen nachzuweisen (wie – beispielsweise – den Gettier-Fällen), dass seine Explikation in obigem Sinne die konventionelle Bewertung dieser Fälle reproduziert. David Chalmers leistet dasselbe in ‚The Components of Content‘ (Chalmers 2002) für verschiedene Gedankenexperimente aus dem Bereich der Semantik propositionaler Einstellungen, Robert Nozick argumentiert in *Philosophical Explanations* (Nozick 1981) für die Adäquatheit seiner closest-continuer theory, indem er verschiedene Gedankenexperimente aus der Debatte um transtemporale personale Identität durchgeht und zeigt, dass seine Theorie in der Anwendung auf Gedankenexperimente aus der Tradition zu den erwünschten Resultaten kommt. Diese Liste könnte man beliebig verlängern. Wichtig ist, dass diese Gedankenexperimente als Adäquatheitskontexte allgemein akzeptiert werden.

Warum nimmt man dazu ausgedachte und seltsame Szenarien anstatt alltägliche Situationen? Nun, manchmal nimmt man auch alltägliche Situationen. Aber erstens ist es nicht immer einfach eine alltägliche Situation zu finden, die genau die entsprechenden Eigenschaften aufweist, auf die man die Aufmerksamkeit lenken will. Zum anderen mag in alltäglichen Situationen diese Eigenschaft weniger hervorstechend vorhanden sein, als in einem konstruierten Fall.

Man einigt sich also bei Gedankenexperimenten als Adäquatheitskontexten darauf, dass jede zukünftige Explikation etwas zu diesem Kontext sagen muss, idealer Weise muss jede zukünftige Explikation bezüglich dieses Kontextes zu den verabredeten Beurteilungen kommen.

Letzteres kann – und jetzt schlagen wir die Brücke wieder zurück zur Psychologie – in Fällen, in denen man gezwungen ist von der verabredeten Beurteilung abzuweichen, aber auch erfüllt werden, indem man den Adäquatheitskontext auf andere Weise wegerklärt. Im Fall von Frank Jacksons Gedankenexperiment mit Mary bestanden solche Strategien darin, dass man versuchte, eine Theorie zu finden, derzufolge es sich nicht um einen eigentlichen Adäquatheitskontext für den fraglichen Gegenstandsbereich der Theorie handelt, etwa weil es in diesem Kontext um *knowing how*, nicht um *knowing that* geht, oder weil es in diesem Kontext nur um einen anderen Aspekt einer bekannten Tatsache geht, also eigentlich um einen Kontext für Theorien kognitiven Gehalts, aber nicht für Theorien metaphysischer Zusammenhänge zwischen Körper und Geist. Konservativität in Bezug auf den Adäquatheitskontext wird dann dadurch hergestellt, dass man die Theorie, die *prima facie* an dem Kontext scheitert, um eine unabhängige aber kompatible Theorie erweitert, die die ursprüngliche Beurteilung wieder herstellt. Für diese Zusatztheorie gelten selbstverständlich ebenfalls Adäquatheitskontexte, deren Erfüllung unabhängig sichergestellt werden muss.

Wenn wir diese Überlegungen auf den „Streit“ zwischen Psychologen und Philosophen übertragen, macht das Verhalten der Philosophen durchaus Sinn. Philosophen versuchen darauf einzuwirken, welche Frage sie von den Psychologen geklärt und ggf. empirisch untersucht wissen wollen. Indem festgestellt wird, dass psychologische Theorien von ‚Bedeutung‘, ‚Wissen‘, etc. „Intuitionen“ verletzen, soll nicht eigentlich behauptet werden, dass man intuitiv die Dinge besser beurteilen kann als derjenige, der empirisch nachgesehen hat, sondern dass die vorgelegte Antwort bzw. das anvisierte Forschungsprogramm an der eigentlich interessierenden Fragestellung vorbeigeht. Das Explicandum ist eben anders intendiert, als die Psychologen meinen.

3. ARBEITSTEILUNG IN DER PHILOSOPHIE

Die moderne Philosophie ist arbeitsteilig organisiert. Wir kennen systematische Teilbereiche, in denen unterschiedliche Theorien miteinander konkurrieren, die auf eben jene Teilbereiche

zugeschnitten sind. Im vergangenen Zeitalter der „Systeme“⁶ konnte sich die interne Kritik von Gedankenexperimenten immer nur auf ihre theoretische Rolle beschränken: Es konnte bestenfalls gezeigt werden, dass (gegeben die Intentionen ihres jeweiligen Autors) die Theorie inkonsistent war, man konnte aber nicht sinnvoll mit Adäquatheitskontexten an sie herantreten, die von ihrem jeweiligen Autor nicht eindeutig abgesegnet waren. Jedes System kann dadurch gegen Kritik immunisiert werden, indem man die problematischen Adäquatheitskontexte von der Theorienbewertung einfach ausschließt. Damit die moderne Form philosophischer Arbeitsteilung möglich ist, einigt man sich über die Diskussion von Gedankenexperimenten auf die relevanten Adäquatheitskontexte, innerhalb derer zunächst Begriffe expliziert werden sollen, die dann ggf. das Grundgerüst empirischer Theorien bilden. Diese Adäquatheitskontexte sind ausschlaggebend dafür, welche Richtung die gemeinsamen Forschungsanstrengungen in Zukunft nehmen sollen, entsprechend kontrovers sind sie auch diskutiert. Diese Diskussionen dürfen gerade nicht als Ausdruck der unüberbrückbaren Unvereinbarkeit verschiedener Positionen der Gegenwartsphilosophie verstanden werden. Über diese „Streitigkeiten“ einigen sich die Wissenschaftler auf das gemeinsame Verständnis des jeweiligen Forschungsprogramms.

Natürlich bemisst sich die Güte eines Gedankenexperiments zunächst daran, ob es mittelbar zu den intendierten Überzeugungsänderungen führt. D. h., ob es mittelbar dazu geeignet ist, die inhaltliche Richtung eines Forschungsprogramms zu beeinflussen. Dies ist sicherlich die Hauptintention desjenigen, der ein Gedankenexperiment vorträgt: Er möchte zum Ausdruck bringen, dass der gegenwärtige Stand der Explikation oder die Richtung, in der nach einer gesucht wird, Kontexte außer Acht lässt, von denen wir alle wollen sollen, dass sie durch die Explikation bedacht werden.

Langfristig ist ein Gedankenexperiment gut, wenn es in den Kanon der dauerhaften Adäquatheitskontexte aufgenommen wird. Gedankenexperimente, die von der Forschergemeinde akzeptiert werden, wandeln sich zu Standards, die die Grundlage der weiteren Arbeit bilden. Das ändert sich nicht dadurch, dass ein Adäquatheitskontext durch eine Zusatztheorie T2 als ein Kontext erwiesen werden kann, der nicht zum Gegenstandsbereich von T1, der eigentlich kritisierten Theorie gehört. Der Adäquatheitskontext bleibt auch dann „virulent“, schließlich kann sich durch die weitere theoretische Entwicklung herausstellen, dass T1 und T2 inkompatibel sind, oder T2 erweist sich aus unabhängigen Gründen als falsch oder tatsächlich nicht auf den fraglichen Kontext zutreffend, etc. Wichtig ist, dass ein solcher Adäquatheitskontext dauerhaft als solcher akzeptiert wird.

Natürlich bedeutet das nicht, dass uns die Welt diese Erklärungswünsche auch erfüllen lässt. Es kann immer sein, dass wir Adäquatheitsbedingungen fordern, die keine wahre Theorie des fraglichen Gegenstandsbereichs je erfüllen können wird, weil wir uns einfach über die Natur der empirischen Phänomene täuschen. Solange man das aber nicht gezeigt hat, darf man darauf insistieren, das erklärt zu bekommen, was man erklärungsbedürftig findet.

Soweit die Charakterisierung von Gedankenexperimenten und ihrer Funktion in Bezug auf Begriffsexplikationen. Für ihren Erfolg kommt es darauf an, dass sie durch die Wissenschaftsgemeinschaft akzeptiert werden. Dafür gelten verschiedene Bedingungen, die aus Platzgründen hier nicht behandelt werden können.⁷

4. WIRD ES AUCH IN DER ZUKUNFT NOCH GEDANKENEXPERIMENTE GEBEN?

Natürlich steht es jedem frei, sich aus dem Methodenarsenal der Wissenschaften soweit zu bedienen, wie es nicht zu pragmatischen Widersprüchen führt (wenn man etwa Methoden

⁶ Dies betrifft eben jene Phasen der vor-analytischen Philosophie, von denen sich die Mitglieder des Wiener Kreises absetzen wollten.

⁷ Diese Bedingungen werden in Häggqvist 1996 und Cohnitz 2006 diskutiert.

verwendet, die nach eigener Auffassung in dieser Verwendung unzulässig sind). Wenn man dann auf Methoden verzichtet, die nachgewiesenermaßen nicht zu solchen Widersprüchen führen, weil man sie aus anderen Gründen ablehnt, ist dies natürlich ebenfalls legitim. Man ist aber dadurch nicht davor geschützt, dass andere Gedankenexperimente in kritischer Absicht an diese Theorien herantragen. Kritische Gedankenexperimente zurückzuweisen mit dem Hinweis darauf, dass man sie nicht mag, ist nicht gerechtfertigt. Man kann Gedankenexperimente zurückweisen, weil sie aus methodologischen Gründen schlecht sind, man kann sie aber nicht nur deswegen zurückweisen, weil sie Gedankenexperimente sind. Das wäre methodologisch genauso falsch, wie empirische Daten zu ignorieren, weil sie gegen die eigene Theorie zu sprechen scheinen. Natürlich gibt es Umstände, in denen man eine Theorie gegen die empirischen Daten in Schutz nehmen muss, und auch Umstände in denen man an einer Theorie festhält, obwohl es Gedankenexperimente gibt, die ihre explikative Adäquatheit in Frage stellen. Solche Umstände bestehen etwa dann, wenn keine befriedigende alternative Theorie existiert, die die akzeptierten Adäquatheitskontexte ansonsten im selben Grad erfüllt. Man kann und wird dann versuchen, die problematischen Adäquatheitskontexte wegzuerklären. Gelingt dies nicht, kann es dennoch gerechtfertigt sein, die Theorie nicht aufzugeben, solange kein adäquater Nachfolger gefunden ist.

Ein etwas anders gelagerter Fall liegt bei der oben rekonstruierten Diskussion zwischen Psychologen und Philosophen vor, in der die Psychologen nicht nachvollziehen können, wieso die Philosophen meinen, mit ihren ausgedachten Experimenten die empirischen Theorien in Frage stellen zu können. Aus der Sicht der Psychologie, sofern sie der Auffassung ist, sie habe bestimmte Begriffe für ihre Theoriebildung gewählt, weil sie für die Erklärungs- bzw. Forschungszwecke besonders fruchtbar schienen, ist es durchaus legitim, die Ansprüche der Philosophen zu ignorieren. Diese Forderungen kann man nur dann *nicht* zurückweisen, wenn man den Anspruch erhebt, an einem gemeinsamen Problem zu arbeiten. Wenn der Psychologe behauptet, er habe etwas zu einem alten philosophischen Problem herausgefunden, muss er sich auch mit Gedankenexperimenten „nerven“ lassen.

Ob es nicht bisweilen heuristisch klüger wäre, den Hinweis der Philosophen ernst zu nehmen, dass man mit dem psychologischen Bedeutungs- oder Wissensbegriff vielleicht an Phänomenen vorbeigeht, die wir intuitiv für erklärungsbedürftig halten, kann man – wie so Vieles – a priori nicht beurteilen.

Ob es in der Zukunft des Philosophie des Geistes auch weiterhin Gedankenexperimente geben wird, kann man auch nicht a priori beantworten. Es gibt aber – denke ich – keine guten Gründe ihren Einfluss zu vermindern. Da Gedankenexperimente wichtige Funktionen in der Methodologie der modernen Philosophie erfüllen, sollte man eher hoffen, dass sie ein zentraler Bestandteil der philosophischen Debatte bleiben. Selbst wenn man einen ähnlich guten Ersatz für sie hätte (wobei ich nicht weiß, wie dieser aussehen könnte), wäre es nicht zuletzt aufgrund ihres Unterhaltungswerts schade, wenn sie aus der Philosophie verschwänden.

Literatur:

- Carnap, Rudolf (1967): *Logical Foundations of Probability*, Routledge.
- Chalmers, David (2002): „The Components of Content“, ders. (Hrsg.): *Philosophy of Mind: Classical and Contemporary Readings*, Oxford University Press, 608-633.
- Cohnitz, Daniel (2006): *Gedankenexperimente in der Philosophie*, Mentis.
- Cohnitz, Daniel (2006a): „Poor Thought Experiments? A Comment on Peijnenburg and Atkinson“, *Journal for General Philosophy of Science* 37, 373-392.
- Degenaar, M. (1996): *Molyneux's Problem: Three Centuries of Discussion on the Perception of Forms*, Kluwer.
- Dretske, Fred (1999): *Knowledge and the Flow of Information*, CSLI Press.

- Gopnik, Alison; Schwitzgebel, Eric (1998): „Whose concepts are they, anyway? The role of philosophical intuition in empirical psychology“, DePaul, M.R.; Ramsey, W. (Hrsg.): *Rethinking Intuition. The Psychology of Intuition and Its Role in Philosophical Inquiry*, Rowman & Littlefield, 75-91.
- Häggqvist, Sören (1996): *Thought Experiments in Philosophy*, Almqvist & Wiksell International.
- Kornblith, Hilary: *Knowledge and Its Place in Nature*, Clarendon Press.
- Machery, Edouard; Mallon, Ron; Nichols, Shaun; Stich, Stephen P.: „Semantics, cross-cultural style“, *Cognition* 92, B1-B12.
- Nozick, Robert (1981): *Philosophical Explanations*, Harvard University Press.
- Peijnenburg, Jeanne; Atkinson, David (2003): „When are thought experiments poor ones?“, *Journal for General Philosophy of Science* 34, 305-322.
- Prinz, Wolfgang (2007): „Philosophie nervt. Eine Polemik“, Spät, Patrick (Hrsg.): *Zur Zukunft der Philosophie des Geistes*, Mentis.