

Ansgar Beckermann

Wie real sind intentionale Zustände? Dennett zwischen Fodor und den Churchlands

1. Quine zufolge unterscheiden sich ontologische Annahmen, d.h. Annahmen darüber, was es wirklich gibt bzw. welche Entitäten tatsächlich real sind, in ihrem Status nicht von anderen wissenschaftlichen Annahmen. Alle Annahmen dieser Art haben synthetischen Charakter. Sie sind nur dadurch zu rechtfertigen, daß sich auf ihrer Grundlage bessere wissenschaftliche Theorien formulieren lassen, d.h. einfachere und umfassendere Theorien, die es ermöglichen, zukünftige Erfahrungen besser vorauszusagen. Dies gilt für die Annahme der Existenz von Steinen, Bäumen und Tieren ebenso wie für die Annahme der Existenz von Protonen, Elektronen und Neutrinos und für die Annahme der Existenz von Klassen und Eigenschaften ebenso wie für die Annahme der Existenz der homerischen Götter Zeus, Hera und Poseidon.

"As an empiricist I continue to think of the conceptual scheme of science as a tool, ultimately, for predicting future experience in the light of past experience. Physical objects are conceptually imported into the situation as convenient intermediaries - not by definition in terms of experience, but simply as irreducible posits comparable, epistemologically, to the gods of Homer." (Quine 1951, S. 44)

Damit ist natürlich nicht gesagt, daß alle Annahmen dieser Art gleich gut begründet sind. Die meisten von uns (Quine eingeschlossen) halten die Annahme, daß es Steine und Bäume gibt, für gut begründet, während sie von der Annahme der Existenz Poseidons und Heras nicht im gleichen Maße überzeugt sind. Aber das ist nicht entscheidend. Entscheidend ist vielmehr, daß wir Homer damit einen wissenschaftlichen und nicht einen begrifflichen bzw. philosophischen Irrtum unterstellen. D.h., vom erkenntnistheoretischen Standpunkt aus unterscheiden sich physikalische Gegenstände und Homers Götter nur dem Grad, aber nicht der Art nach. Dasselbe gilt für die ontologischen Annahmen, die wir machen, wenn wir Gegenstände auf der Ebene der Atome und Elementarteilchen postulieren, um das Verhalten makroskopischer Gegenstände zu erklären. Oder wenn wir annehmen, daß das Geschehen auf der Ebene der Elementarteilchen durch eine endliche Anzahl von Kräften bestimmt ist. Selbst die abstrakten Gegenstände der Mathematik - Klassen und Klassen von Klassen - befinden sich erkenntnistheoretisch auf einer Stufe mit den physikalischen Gegenständen und den homerischen Göttern. Sie unterscheiden sich nur in dem Maße, in dem sie dazu beitragen, unsere Erfahrungen in einen systematischen Zusammenhang zu bringen. Je größer der erklärende Wert einer ontologischen Annahme ist, umso stärker ist sie in unserem Überzeugungssystem verankert. Und da dies für alle wissenschaftlichen Aussagen gilt, unterscheiden sich ontologische Annahmen in

ihrem erkenntnistheoretischen Status nicht von anderen wissenschaftlichen Hypothesen.

Für die gegenwärtige Debatte um die Realität intentionaler Zustände ist diese Grundposition Quines von besonderer Bedeutung, weil alle Teilnehmer an dieser Debatte diese Position insofern teilen, als sie der Meinung sind, daß intentionale Zustände *nur* dann real sein können, wenn sich zeigen läßt, daß die Annahme solcher Zustände für die adäquate Erklärung menschlichen Verhaltens unentbehrlich ist.

Ganz im Sinne Quines argumentiert z.B. Jerry Fodor, das entscheidende Argument für die Annahme der Existenz intentionaler Zustände liege in der Tatsache, daß die Theorie, in der intentionale Zustände ihren natürlichen Platz haben, nämlich die intentionale Alltagspsychologie ('commonsense belief/ desire-psychology'), eine erfolgreiche, keineswegs oberflächliche und letztlich sogar unentbehrliche Theorie zur Erklärung und Voraussage menschlichen Verhaltens sei. Schon einfache Beispiele belegen Fodor zufolge diese Auffassung. Jemand ruft mich an und fragt, ob ich am nächsten Dienstag in Berlin einen Vortrag halten könne. Vielleicht ist das einzige, was ich darauf antworte: "Ja, vielen Dank. Ich werde um 12.15 Uhr am Bahnhof Zoo sein". In diesem Fall wäre die Äußerung dieser Wörter also das einzige beobachtbare Verhalten, auf das sich mein Gesprächspartner bei der Voraussage meines Verhaltens stützen kann. Aber das genügt auch. Den ganzen restlichen Teil der Aufgabe, mein weiteres Verhalten vorherzusagen, d.h. die Lücke zwischen Äußerung und zukünftigem Verhalten zu schließen, erledigt die intentionale Alltagspsychologie. Und diese Theorie funktioniert so gut, daß sich mein Gesprächspartner mit ziemlicher Sicherheit darauf verlassen kann, daß ich am nächsten Dienstag um 12.15 Uhr am Bahnhof Zoo sein werde. Wenn ich nicht da sein sollte, wird das jedenfalls eher an einer Verspätung des Zuges liegen als an einem generellen Fehler der Alltagspsychologie, auf die er sich bei seiner Voraussage stützt.

Voraussagen dieser Art sind möglich, weil die intentionale Alltagspsychologie uns zum Beispiel sagt, wie wir aus den Wörtern, die jemand äußert, auf seine Absichten schließen können (wenn jemand, der Deutsch spricht, die Wörter "Ich werde um 12.15 Uhr am Bahnhof Zoo sein" äußert, dann beabsichtigt er in der Regel um 12.15 Uhr am Bahnhof Zoo zu sein) und wie wir aus seinen Absichten auf sein Verhalten schließen können (wenn jemand beabsichtigt um 12.15 Uhr am Bahnhof Zoo zu sein, dann wird er in der Regel etwas tun, das - sofern nicht unvorhergesehene Zwischenfälle eintreten - zur Folge haben wird, daß er um 12.15 Uhr am Bahnhof Zoo ist). Und dies alles funktioniert nicht nur bei uns bekannten Menschen, mit deren Psychologie wir sehr gut vertraut sind. Es funktioniert auch bei wildfremden Menschen, die wir noch nie gesehen haben. Und es funktioniert nicht nur unter Laborbedingungen, sondern auch unter den kaum je vollständig überschaubaren Bedingungen des normalen Lebens. Dies ist ein sehr bemerkenswerter Erfolg, der Fodor zu dem Kommentar veranlaßt: Wenn wir das Wetter auch nur halb so gut voraussagen könnten, würde niemand von uns je nasse Füße bekommen.

Zumindest auf den ersten Blick ist der explanatorische und prognostische Erfolg der intentionalen Alltagspsychologie also beeindruckend. Dennoch ist Fo-

dors Argumentation nicht unwidersprochen geblieben. P.M. Churchland z.B. hat sehr nachdrücklich die These vertreten, die Einschätzung, die intentionale Alltagspsychologie sei eine erfolgreiche und erklärungskräftige Theorie, beruhe auf einer sehr vordergründigen Bewertung der Tatsachen und eine genauere Analyse führe zu ganz anderen Ergebnissen.¹ Denn erstens, so Churchland, müsse man nicht nur die Erfolge, sondern auch die Mißerfolge der Alltagspsychologie berücksichtigen, d.h. besonders das Ausmaß und die Tragweite dieser Mißerfolge. Zweitens sei es wichtig, die Geschichte der Alltagspsychologie mitzuberücksichtigen, und d.h. insbesondere ihre Fähigkeit, sich weiter zu entwickeln, ihre Fruchtbarkeit und die von ihr zu erwartenden zukünftigen Erfolge. Drittens schließlich müsse man bedenken, welche Theorien zur Erklärung von Verhalten vor dem Hintergrund dessen, was wir in den letzten Jahrzehnten über uns selbst als biologische Wesen gelernt haben, wahrscheinlich wahr sind. D.h. man müsse auch die Frage stellen, wie sich die Alltagspsychologie zu fruchtbaren und gut etablierten Theorien in Nachbarbereichen verhält - wie etwa der Evolutionstheorie oder der Neurobiologie. Im Hinblick auf alle drei Punkte schneidet Churchland zufolge die Alltagspsychologie schlecht ab.

2. Mein Ziel in diesem Aufsatz ist allerdings nicht, in dem von Fodor und Churchland geführten Streit den Schiedsrichter zu spielen. Die Darstellung der Argumente dieser beiden Autoren hatte nur den Zweck, einen Argumenttyp noch deutlicher werden zu lassen, mit dem vor dem Hintergrund der Quineschen Grundposition in den letzten Jahren um die Frage der Realität intentionaler Zustände gestritten worden ist. Im folgenden möchte ich vielmehr die Position Dennetts etwas genauer untersuchen, die in diesem Zusammenhang auf den ersten Blick irritierend erscheinen muß. Denn Dennett betont auf der einen Seite - ebenso wie Fodor - den prognostischen und explanatorischen Erfolg und letztlich sogar die Unentbehrlichkeit der intentionalen Alltagspsychologie.² Auf der anderen Seite weigert er sich jedoch, Fodors Schluß auf die uneingeschränkte Realität intentionaler Zustände mitzumachen. Denn trotz ihrer Nützlichkeit - und vielleicht sogar Unentbehrlichkeit - bei der Erklärung und Prognose menschlichen Verhaltens haben intentionale Zustände Dennett zufolge nur den Status theoretischer Konstrukte (*abstracta*), die nicht im strengen Sinne als real angesehen werden können. Welche Gründe sprechen Dennett zufolge für diese, auf den ersten Blick doch eher unverständliche Zwischenposition? Bevor ich mich dieser Frage zuwende, möchte ich zuerst allerdings noch kurz auf Dennetts Argumente für den Erfolg bzw. die Unentbehrlichkeit der intentionalen Psychologie eingehen.

Bei der Entwicklung seiner Theorie intentionaler Systeme geht Dennett bekanntlich von der Grundthese aus, daß man bei der Erklärung des Verhaltens komplexer Systeme drei grundsätzlich verschiedene Einstellungen einnehmen kann: die *physikalische* Einstellung (*physical stance*), die *funktionale* Einstellung (*design stance*) und die *intentionale* Einstellung (*intentional stance*). In physikalischer Einstellung erklärt man das Verhalten eines Systems, indem man sich klarmacht, aus welchen physikalischen Komponenten das System besteht, wie diese Komponenten aufgrund der für sie geltenden physikalischen Naturgesetze

interagieren und welches Systemverhalten aus diesen Interaktionen resultiert. In funktionaler Einstellung überlegt man sich, welche funktionalen Bestandteile das System enthält und welches Systemverhalten sich ergibt, wenn alle diese Bestandteile so funktionieren, wie sie funktionieren sollen. Und in intentionaler Einstellung schließlich erklärt man das Verhalten eines Systems, wenn man unterstellt, daß das System über bestimmte Informationen verfügt (bestimmte Überzeugungen hat), daß es bestimmte Ziele verfolgt und daß es sich angesichts dieser Informationen und Ziele rational verhält, d.h. das tut, was unter der Voraussetzung, daß die Informationen zutreffen, tatsächlich zur Realisierung seiner Ziele führt.

Besonders in Dennetts frühesten Aufsätzen wird jedoch nicht recht deutlich, welche Motive eigentlich dafür verantwortlich sind, daß man bei dem Versuch, das Verhalten eines Systems zu erklären, in dem einen Fall die eine und in einem anderen Fall eine andere Einstellung wählt. In (1971) erhält man den Eindruck, daß diese Wahl im wesentlichen von Praktikabilitäts Gesichtspunkten bestimmt wird. Wir wählen die jeweils 'höhere' Einstellung, wenn wir Erklärungen auf einer niedrigeren Stufe nicht erreichen können, da uns die relevanten Informationen fehlen oder der Rechenaufwand wegen der Komplexität der physikalischen oder funktionalen Strukturen zu groß ist.

"One seldom adopts the physical stance in dealing with a computer just because the number of critical variables in the physical constitution of a computer would overwhelm the most prodigious calculator." (1971, S. 5)

"The best chess-playing computers these days are practically inaccessible to prediction from either the design stance or the physical stance; they have become too complex for even their own designers to view from the design stance. A man's best hope of defeating such a machine in a chess match is to predict its responses by figuring out as best he can what the best or most rational move would be, given the rules and goals of chess. That is, one assumes not only (1) that the machine will function as designed, but (2) that the design is optimal as well, that the computer will 'choose' the most rational move." (ebd.)

Wenn das so wäre, wären intentionale Erklärungen aber höchstens in einem praktischen Sinne unentbehrlich. Prinzipiell ließe sich jedes Verhalten auch von der funktionalen oder physikalischen Einstellung aus erklären. In einigen Fällen wären *uns* diese Erklärungen nur nicht zugänglich. Und prinzipiell wären die niederstufigen Erklärungen sogar die besseren, weil genaueren Erklärungen, da sie uns nicht nur sagen, wie sich das System verhielte, wenn es so funktionierte, wie es funktionieren sollte, bzw. wenn es optimal konstruiert wäre, sondern wie sich das System unter den gegebenen Bedingungen tatsächlich verhält.

Schon in (1981a) hat Dennett jedoch eine alternative Begründung entwickelt, derzufolge intentionale Erklärungen auch in einem anderen Sinne unentbehrlich sind. Diese Begründung beruht auf der These, daß jedem, der bestimmten Systemen gegenüber die intentionale Einstellung *nicht* einnimmt, bestimmte Regelmäßigkeiten (Muster) im Verhalten dieser Systeme unzugänglich werden. Denn diese Regelmäßigkeiten oder Muster sind sozusagen nur von der intentionalen Einstellung her 'sichtbar'.

Dennett erläutert diese These mit der folgenden Überlegung. Angenommen, auf dem Mars gäbe es außergewöhnliche intelligente Wesen, die eines Tages unsere Erde besuchen. Daß sie außergewöhnlich intelligent sind, soll z.B. heißen, daß wir für sie nicht schwerer zu verstehen sind als ein Thermostat für einen einigermaßen begabten Ingenieur. Für diese hyperintelligenten Besucher vom Mars gibt es daher keinen praktischen Grund, die intentionale Einstellung einzunehmen. Sie können - gleichsam als Laplacesche Superphysiker - z.B. das gesamte Geschehen an der Wall Street schon auf der Ebene der Mikrophysik vollständig beschreiben und vorhersagen. Sie können das gesamte Verhalten aller beteiligten Körper prognostizieren, ohne auch nur einen von ihnen als intentionales System auffassen zu müssen.

Dennett zufolge hätte das Verfahren unserer Besucher vom Mars, alle Bewegungen auf der Ebene der Mikrophysik zu analysieren und vorauszusagen, jedoch einen entscheidenden Nachteil:

"Our imagined Martians might be able to predict the future of the human race by Laplacean methods, but if they did not also see us as intentional systems, they would be missing something perfectly objective: the *patterns* in human behavior that are describable from the intentional stance, and only from the intentional stance, and that support generalizations and predictions." (1981a, S. 25)

Denken wir etwa an das Verhalten eines Börsenmaklers, der 500 Aktien von General Motors kaufen möchte. Auf der Grundlage ihrer umfassenden mikrophysikalischen Kenntnisse könnten die Besucher vom Mars die Bewegungen seiner Finger beim Wählen der Telefonnummer ebenso präzise vorhersagen wie die Schwingungen seiner Stimmbänder beim Formulieren des Kaufauftrags. Aber sie würden dabei übersehen, daß unzählig viele andere Muster von Fingerbewegungen und Stimmbänderschwingungen für den Ablauf des Börsengeschehens genau denselben Effekt gehabt hätten. D.h., sie würden ein in der Welt real existierendes Muster nicht erkennen. So wie jemand die Funktionsweise eines Verbrennungsmotors nicht verstanden hat, wenn er nicht sieht, daß viele verschiedene Arten von Gegenständen die Funktion einer Zündkerze ausüben können, haben die Besucher vom Mars das Börsengeschehen nicht verstanden, wenn sie nicht sehen, daß es viele verschiedene Arten gibt, eine Kauforder für Aktien zu plazieren.

Ich denke, daß Dennetts Argumentation hier auf nichts anderes hinausläuft als auf die bekannte Unterscheidung von Körperbewegungen und Handlungen. Und in der Tat hat er sicher recht, wenn er noch einmal unterstreicht, daß es neben der Möglichkeit, die Körperbewegungen, die das Verhalten eines Systems ausmachen, aufgrund von physischen Merkmalen zu klassifizieren, auch noch andere Möglichkeiten der Klassifikation gibt und daß sich auch für diese (anders individuierten) Klassen von Bewegungen interessante Gesetzmäßigkeiten feststellen lassen. Allerdings macht er nicht ganz klar, warum wir diese Gesetzmäßigkeiten nur von der intentionalen Einstellung her sehen können. Die intentionale Einstellung ist durch eine bestimmte Art der Erklärung von Verhalten charakterisiert und nicht durch den Typ des Verhaltens, das erklärt wird. Unter der Hand scheint Dennett hier seine Ausgangsposition, daß man das Verhalten komplexer Systeme immer auf allen drei Ebenen - der physischen, der funktionalen,

und der intentionalen - erklären könne, zu modifizieren und anzudeuten, daß dies nicht unabhängig davon gilt, wie das zu erklärende Verhalten klassifiziert und beschrieben wird. Verhalten im Sinne von Handeln scheint auch Dennett zufolge nur auf der intentionalen Ebene erklärbar zu sein. Allerdings gibt er für diese Auffassung kaum Gründe an.

3. Aber wie dem auch sei, kommen wir zurück zu der Frage, welche Gründe Dennett dazu bewegen, trotz des enormen prognostischen Erfolgs der intentionalen Alltagspsychologie und trotz der Tatsache, daß diese Theorie auch seiner Meinung nach in gewissem Sinne unentbehrlich ist, den Fodorschen Schluß auf die uneingeschränkte Realität intentionaler Zustände nicht mitzumachen.

In (1987f) betont Dennett - wie an vielen anderen Stellen auch - den Zusammenhang seiner Überlegungen mit der Position Quines, der auf der einen Seite die Auffassung vertritt, daß es intentionale Zustände in einem strengen Sinne eigentlich nicht gebe, auf der anderen Seite aber zugleich betont, daß seiner Meinung nach das intentionale Vokabular als Teil einer "wesentlich dramatischen" Ausdruckform "praktisch unverzichtbar" sei (Quine 1960, S. 219, 221). Zum ersten Teil dieser Position kommt Quine aufgrund seiner Überlegungen zur Unbestimmtheit der Übersetzung, die letzten Endes in der These gipfeln, daß es keine Frage der Tatsachen ist, ob ein bestimmter sprachlicher Ausdruck *A* oder *B* bedeutet, und daß es aus denselben Gründen auch keine Frage der Tatsachen ist, ob ein Fremder, dessen Überzeugungen man untersucht, eher *A* als *B* glaubt.

"The metaphor of the black box, often so useful, can be misleading here. The problem is not one of hidden facts, such as might be uncovered by learning more about the brain physiology of thought processes. To expect a distinctive physical mechanism behind every genuinely distinct mental state is one thing; to expect a distinctive mechanism for every purported distinction that can be phrased in traditional mentalistic language is another. The question whether [...] the foreigner *really* believes *A* or believes rather *B*, is a question whose very significance I would put to doubt. This is what I am getting at in arguing the indeterminacy of translation." (Quine 1970, S. 180 f.)

Welche genaue Bedeutung ein bestimmter sprachlicher Ausdruck hat bzw. was genau jemand glaubt oder wünscht, das sind Quine zufolge also Fragen, deren Beantwortung nicht allein von Tatsachen abhängt. Im Hinblick auf diese Fragen gibt es daher immer einen Interpretationsspielraum, der durch normative Prinzipien wie das 'principle of charity' oder projektive Prinzipien ausgefüllt wird wie das Prinzip, daß wir Fremden in der Regel die Überzeugungen zuschreiben, die wir auch selbst unter den gegebenen Umständen hätten. Eine ähnliche Position vertritt auch Putnam, wenn er sich in (1983) kritisch zu den Grundannahmen mancher kognitiver Psychologen und damit auch zum Fodorschen Programm einer Repräsentationalen Theorie des Geistes äußert:

"Belief-desire explanation belongs to the level of what I've been calling *interpretation theory*. It is as holistic and interest-relative as all interpretation. Psychologists often speak as though there were *concepts* in the *brain*. The point of my argument [...] is that there may be *sentence-analogues* and *predicate-analo-*

gues in the brain, but not concepts. 'Mental representations' require interpretation just as much as any other signs do." (S. 154)

Doch dies ist *nicht* die Argumentation Dennetts.³ Denn die These, daß der *Inhalt* intentionaler Zustände durch die objektiven Tatsachen nicht eindeutig determiniert ist und daß es daher bei der Bestimmung dieses Inhalts immer Interpretationsspielräume gibt, ist mit der Annahme der *Realität* intentionaler Zustände durchaus vereinbar. Das Bestehen solcher Unbestimmtheiten und Interpretationsspielräume kann nämlich durch die Annahme erklärt werden, daß Inhalte im Zusammenhang mit intentionalen Zuständen eine ähnliche Rolle spielen wie Zahlen im Zusammenhang mit den quantitativen Größen der Physik. Aus der Tatsache, daß wir Temperaturen in Grad Celsius oder in Grad Fahrenheit und daß wir Längen in Metern oder in Yards angeben können, würde schließlich auch niemand folgern, daß die Temperaturen von Gasen oder die Längen von Stäben keine realen Eigenschaften sind. Dies ist genau die Position, die Davidson in (1989) explizit formuliert:

"We know there is no contradiction between the temperature of the air being 32 fahrenheit or 0 celsius; there is nothing in this 'relativism' to show that the properties being measured are not 'real' [...] in the light of the considerations put forward here, this [i.e., that either of two different interpretations might correctly be put on the same thought (or utterance) of a person] comes to no more than the recognition that more than one set of one person's utterances might be equally successful in capturing the contents of some else's thoughts and speech. Just as numbers can capture all the empirically significant relations among weights or temperatures in infinitely many different ways, so one person's utterance can capture all the significant features of another person's thoughts and speech in different ways. This fact does not challenge the 'reality' of the attitudes or meanings thus variously reported." (S. 16)

Aber wie gesagt, dies ist nicht die Position Dennetts. Davidsons Abschwächung des Realismus geht ihm nicht weit genug.⁴ Wenn sich Dennett auf Quine beruft, dann denkt er nicht in erster Linie an eine einfache Übertragung der Unbestimmtheitstheese auf die Frage nach dem Inhalt von intentionalen Zuständen. Sein Punkt ist vielmehr, daß man im Zusammenhang mit der Frage nach der Realität intentionaler Zustände dieselbe *Argumentationsstrategie* anwenden kann, auf der Quines Überlegungen zur Unbestimmtheitstheese beruhen.⁵ So wie Quine seine Unbestimmtheitstheese mit dem Argument begründet, daß es immer sehr unterschiedliche Übersetzungshandbücher gibt, die mit allen Sprachdispositionen der jeweils untersuchten Sprachgemeinschaft in Einklang stehen, kann man Dennett zufolge die These, daß es keine Frage der Tatsachen ist, welche Überzeugungen und Wünsche eine Person hat, dadurch untermauern, daß man zeigt, daß es durchaus möglich ist, daß zwei sehr unterschiedliche Annahmen über die intentionalen Zustände einer Person mit allen ihren Verhaltensdispositionen gleich gut übereinstimmen.

"How, then, does this make me less a realist than Davidson? I see that there could be two different systems of belief attribution to an individual which differed *substantially* in what they attributed - even in yielding substantially different predictions of the individual's future behavior - and yet where no deeper fact

could establish that one was a description of the individual's *real* beliefs and the other not." (1991, S. 49)

Ein entscheidender Grund für diese Position Dennetts liegt darin, daß sich seiner Meinung nach keine Person jemals völlig rational verhält und daß es deshalb niemals möglich ist, ihr Verhalten restlos auf ein bestimmtes Muster von Wünschen, Präferenzen und Überzeugungen zurückzuführen. Es gibt immer Verhaltensweisen, die nicht zu diesem Muster passen und die daher nur durch die Bezugnahme auf (möglicherweise nicht systematisch vorhersagbare) Störfaktoren erklärt werden können. Selbst wenn zwei Personen, Hans und Paul, eine dritte, Angelika, lange Zeit beobachtet haben, kann es deshalb sein, daß sie zur Erklärung des Verhaltens von Angelika intentionale Interpretationen entwickeln, die sich zwar in den groben Umrissen gleichen, die aber im Detail so deutlich voneinander abweichen, daß Hans aufgrund seiner Interpretation zu ganz anderen Voraussagen kommt als Paul.

Daß es für Angelikas Verhalten zwei unterschiedliche intentionale Interpretationen gibt, heißt natürlich noch nicht, daß diese Interpretationen auch gleich gut sind. Hans und Paul könnten deshalb versuchen, ihre Interpretationen zu vergleichen, indem sie sie am Erfolg ihrer Voraussagen messen. Sie könnten z.B. darauf wetten, was Angelika als nächstes tun wird. Wenn sie das tun, wird es auf der einen Seite sicher viele Situationen geben, bei denen sie übereinstimmen und beide recht haben. Auf der anderen Seite wird es in der Regel aber auch eine Reihe von Situationen geben, bei denen sie übereinstimmen und beide unrecht haben, bei denen sie das Verhalten von Angelika also beide falsch voraussagen. In solchen Fällen werden Hans und Paul beide Angelikas Verhalten auf nicht vorhersagbare Störfaktoren zurückführen. Schließlich wird es aber auch Situationen geben, bei denen Hans und Paul verschiedener Meinung sind, d.h. bei denen sich die Voraussagen ihrer intentionalen Interpretationen unterscheiden, und in diesen Situationen wird mal der eine, mal der andere gewinnen. (Auch bei diesen Gelegenheiten können sie jedoch beide falsch liegen.) Wenn Hans richtig liegt und Paul verliert, scheint das zunächst dafür zu sprechen, daß die intentionale Interpretation von Hans besser, angemessener oder richtiger ist als die von Paul. Aber es ist durchaus möglich, daß beide mal gewinnen und mal verlieren, daß sich in den jeweiligen Gewinnen und Verlusten kein festes Muster abzeichnet und daß daher beide nicht sehen, wie sie ihre jeweilige Interpretation verbessern könnten. Daß Hans manchmal gewinnt, ist in diesem Fall kein Grund für die Annahme, daß seine Interpretation näher an der Wahrheit ist als die Pauls.

Sicher, man könnte denken, daß letzten Endes doch eine der beiden Interpretationen die Oberhand gewinnen wird und daß es nur eine Frage der Zeit ist, wann sich herausstellt, welches die bessere Interpretation ist. Aber, so Dennett: Der zentrale Punkt ist gerade, daß dies nicht so sein muß. Es könnte durchaus sein, daß Hans' und Pauls intentionale Interpretationen dauerhaft nebeneinander bestehen und daß es auch 'in the long run' keinen Grund gibt, die eine der anderen vorzuziehen.

Hierin liegt Dennett zufolge die Analogie zur Quineschen Argumentation für die Unbestimmtheitsthese. Es ist durchaus möglich, daß es verschiedene inten-

tionale Interpretationen einer Person gibt, die mit allen Verhaltensdispositionen dieser Person gleichgut in Einklang stehen. Zumindest in einem solchen Fall ist es keine Frage der Tatsachen, welche Überzeugungen und Wünsche diese Person hat.

4. In (1991) erläutert Dennett diese These zusätzlich an einem abstrakteren Beispiel, das direkt nichts mit intentionalen Interpretationen und der Erklärung menschlichen Handelns zu tun hat. Ausgangspunkt sind dabei die in Abb. 1 gezeigten sechs Verteilungen von schwarzen und weißen Punkten. In einem gewissen Sinn, so Dennett, stellen alle diese Verteilungen verschiedene Punktmuster dar. Es gibt keine zwei Verteilungen, die Punkt für Punkt identisch sind.

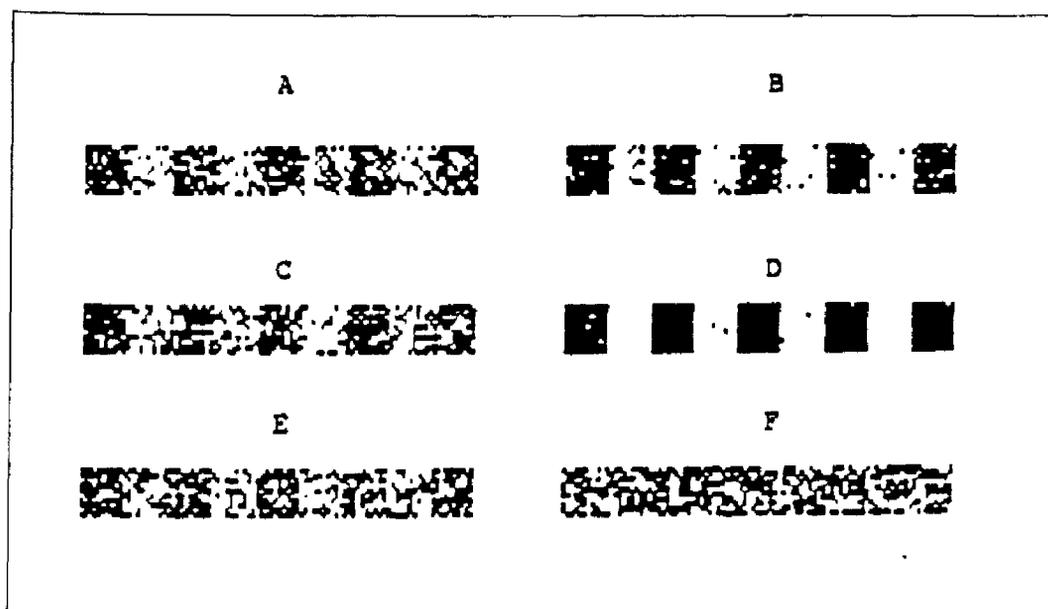


Abbildung 1

Von einem anderen Gesichtspunkt aus betrachtet, zeigen die angeführten sechs Verteilungen aber alle dasselbe Muster. Denn alle beruhen auf demselben Grundprozeß, dem Druck von zehn Zeilen von je neunzig Punkten, wobei zunächst zehn schwarze Punkte gedruckt werden, dann zehn weiße, dann wieder zehn schwarze, usw. Auf diese Weise entstehen im Prinzip fünf gleichgroße schwarze Quadrate, die durch ebenso große weiße Quadrate voneinander getrennt sind. (Dieses Muster nennt Dennett "Balken Code" ("bar code").) Der Unterschied zwischen den Verteilungen A-F besteht nur darin, daß bei ihrer Herstellung das Drucken des "Balken Codes" in unterschiedlichem Maße durch zufallsbedingtes "Rauschen" verzerrt wurde. Der Rauschanteil variierte dabei folgendermaßen:

A: 25%	B: 10%
C: 25%	D: 1%
E: 33%	F: 50%

Mit Bezug auf dieses Beispiel versucht Dennett zunächst noch einmal genauer zu fassen, was es heißt, daß ein Muster oder eine Regelmäßigkeit

"objektiv" ist. Wenn wir uns nur auf unsere Augen verlassen würden, würden wir wahrscheinlich sagen, daß das Balkenmuster in den Verteilungen B und D klar hervortritt, daß es in den Verteilungen A und C noch einigermaßen und in der Verteilung E noch schwach zu erkennen ist, daß sich in der Verteilung F aber keinerlei Muster mehr ausmachen läßt. Doch bei der Frage, welche Muster real gegeben sind, können wir uns nicht ohne weiteres auf unsere Sinnesorgane verlassen. Welche Muster von den Sinnesorganen erkannt werden, das hängt ebenso von ihrer Konstruktion wie vom Training des Beobachters ab. Empirisch ist z.B. erwiesen, daß in einem Gegenstand objektiv vorhandene Symmetrien von unserem visuellen System nur dann einigermaßen verläßlich erkannt werden, wenn die Symmetrieachse parallel zu unserer Körperachse liegt. Auf der Grundlage informationstheoretischer Überlegungen versucht Dennett deshalb ein objektives Kriterium zu formulieren: In einer Punkteverteilung ist ein Muster genau dann real, wenn es unter Bezugnahme auf dieses Muster möglich ist, die Verteilung effizienter zu beschreiben als durch die Aufzählung aller Einzelpunkte. (Diese Aufzählung nennt Dennett die "bit map".) In diesem Sinne ist in der Verteilung D das Balkenmuster real. Denn:

"Frame D [...] can be described as 'ten rows of ninety: ten black followed by ten white, etc., with the following exceptions: dots 57, 88, ...' This expression, suitably encoded, is much shorter than 900 bits long." (1991, S. 32 f.)

So gesehen ist es aber durchaus möglich, vielleicht sogar wahrscheinlich, daß einunddieselbe Punkteverteilung als Realisierung ganz verschiedener Muster angesehen werden kann und daß es nichts gibt, das dafür spricht, daß eines dieser Muster 'realer' ist als ein anderes. Und dies gilt auch allgemein: Wenn zwei Personen dieselben Daten analysieren sollen, dann kann es sein, daß sie in diesen Daten ganz verschiedene Muster realisiert sehen. Dies muß jedoch nicht bedeuten, daß die eine recht hat und die andere unrecht.

"If Jones sees pattern α (with $n\%$ noise) and Brown sees pattern β (with $m\%$ noise) there may be no ground for determining that one of them is right and the other wrong. Suppose they are both using their patterns to bet on the next datum in the series. Jones bets according to the 'pure' pattern α , but budgets for $n\%$ errors when he looks for the odds. Brown does likewise, using pattern β . If both patterns are real, they will both get rich. That is to say, so long as they use their expectation of deviations from the 'ideal' to temper their odds policy, they will do better than chance - perhaps very much better." (1991, S. 35)

Einige Seiten später (1991, S. 42 ff.) gibt Dennett dieser Argumentation jedoch noch einmal eine andere, interessante Wende. Im Abschnitt "The Reality of Intentional Patterns" betont Dennett zunächst noch einmal, daß er in allem, was den prognostischen Erfolg der intentionalen Psychologie angehe, mit Fodor völlig einer Meinung sei. Unterschiedliche Auffassungen gebe es zwischen Fodor und ihm nur in Bezug auf die Frage, wie man diesen Erfolg erklären könne und was aus einer solchen Erklärung über die Realität intentionaler Zustände folge.

Fodor, den Dennett als Superrealisten ("industrial-strength Realist") bezeichnet, sei der Meinung, daß intentionale Zustände nur real sein können, wenn das Muster intentionaler Zustände, das wir einer Person zur Erklärung ihres Verhaltens zuschreiben, durch ein korrespondierendes Muster von Gehirnzuständen

realisiert sei. Wenn wir das Verhalten einer Person (das uns aus der Perspektive der intentionalen Alltagspsychologie als ein Muster mit Rauschen erscheine) nur teilweise unter Bezugnahme auf die intentionalen Zustände dieser Person erklären können, müsse das daher daran liegen, daß dieses Verhalten tatsächlich auf zwei sich überlagernden Prozessen beruhe: einem Grundprozeß, in dem das Muster der intentionalen Psychologie unverfälscht realisiert sei, und einem Störprozeß (der z.B. auf Wahrnehmungs- und Konzentrationsstörungen beruhen könne), der das Resultat des Grundprozesses nachträglich verzerrt.⁶

Wenn das so wäre, entspräche der Mechanismus, der dem Verhalten von Personen zugrundeliegt, offenbar genau dem Mechanismus, der zu den in der Abb. 1 gezeigten Verteilungen von schwarzen und weißen Punkten geführt hat. Auch bei diesem Mechanismus wird zunächst ein reines Balkenmuster erzeugt, das nachträglich durch Störeinflüsse verzerrt wird. Für Fodor ist dies die einzig plausible Annahme, mit der wir den Erfolg der intentionalen Alltagspsychologie erklären können. Aber er gesteht selbst zu, daß diese Argumentation sicher keinen a priori Beweis darstellt.

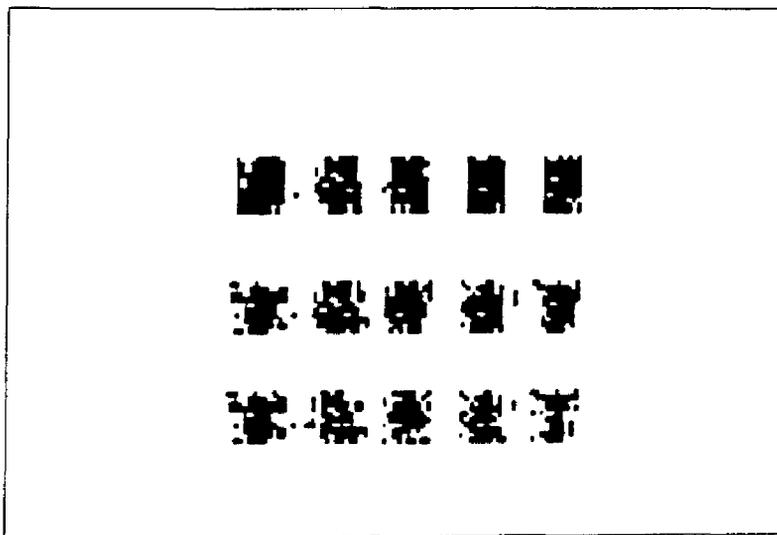


Abbildung 2

Dennett greift genau diesen letzten Punkt auf, indem er - wieder anhand von Verteilungen von schwarzen und weißen Punkten - zu zeigen versucht, daß Muster, die wie Balkenmuster mit ein wenig Rauschen aussehen, auch durch Prozesse erzeugt werden können, die mit der Erzeugung von Balkenmustern zunächst einmal gar nichts zu tun haben. Zu diesem Zweck führt er die in Abb. 2 gezeigten Punkteverteilungen an, über deren Entstehung er folgendes schreibt.

"[...] the top frame shows a pattern created by a normal distribution of black dots around means at $x = 10, 30, 50, 70$ and 90 [...]; the middle and bottom frames were created by successive applications of a very simple contrast enhancer applied to the top frame: a vertical slit 'window' three pixels high is thrown randomly onto the frame; the pixels in the window vote, and majority rules." (1991, S. 44)

Dennett zufolge ist es natürlich möglich, die Punkteverteilungen in der Abb. 2 von den Punkteverteilungen in der Abb. 1 zu unterscheiden. Aber dies ist für ihn nicht entscheidend.

"My point is that *even if* the evidence is substantial that the discernible pattern is produced by one process rather than another, it can be rational to ignore those differences and use the simplest pattern description (e.g. *bar code*) as one's way of organizing that data." (1991, S. 44)

D.h., selbst wenn wir wissen, daß es in dem Prozeß, auf den eine bestimmte Punkteverteilung zurückgeht, keinen Teilprozeß gibt, der direkt als ein Prozeß zur Erzeugung von Balken Code verstanden werden kann, kann es sinnvoll sein, das erzeugte Muster als Balken Code zu deuten und so zu tun, als ginge dieses Muster doch auf einen solchen Prozeß zurück.

5. Ich denke, daß in dieser Argumentation ein zentrales Motiv der Dennettschen Überlegung deutlich wird, ein Motiv, das sich schon in seinen frühesten Aufsätzen findet: die Unterscheidung von externen und internen Verhaltensklärungen.⁷ Und in dieser Unterscheidung liegt meiner Meinung nach auch der entscheidende Schlüssel für die Antwort auf die Frage, warum Dennett trotz des großen explanatorischen und prognostischen Erfolgs der intentionalen Alltagspsychologie intentionalen Zustände nicht den Status uneingeschränkter Realität zuerkennen möchte.

Grundsätzlich läßt sich die Situation, um die es hier geht, folgendermaßen darstellen. Nehmen wir an, daß wir es mit einem System S zu tun haben, das über ein etwas umfangreicheres Verhaltensrepertoire verfügt. Wenn wir dieses System eine Zeitlang beobachten, dann werden wir in seinem Verhalten wahrscheinlich einige Regelmäßigkeiten feststellen. D.h. genauer gesagt, dann werden wir feststellen, daß sich für das Verhalten von S einige interessante gesetzesartige Aussagen formulieren lassen, wenn wir dieses Verhalten mit Hilfe eines geeigneten Begriffssystems klassifizieren. Auf der Grundlage unserer Beobachtungen können wir mit Hilfe dieses Begriffssystems im Laufe der Zeit also eine Theorie T_S zur Erklärung und Vorhersage des Verhaltens von S entwickeln. Ich nenne eine solche Theorie eine externe Verhaltenstheorie, da sie ohne direkte Kenntnis der inneren Struktur von S entwickelt wird.

Je nachdem, wie komplex das Verhalten von S ist, wird auch T_S mehr oder weniger komplex ausfallen. Im einfachsten Fall wird T_S aus einer Reihe von gesetzesartigen Aussagen der folgenden Art bestehen:

- (1) In Situationen vom Typ A verhält sich S (in der Regel) in der Weise X .
- In Situationen vom Typ B verhält sich S (in der Regel) in der Weise Y .
- usw.

Wenn das Verhalten von S komplexer ist, d.h. insbesondere, wenn S sich in Situationen desselben Typs bei verschiedenen Gelegenheiten verschieden verhält, reichen Aussagen dieser Art aber nicht mehr aus. In diesem Fall ist es un-

umgänglich anzunehmen, daß S verschiedene theoretisch-funktionale Zustände annehmen kann, Zustände, die für das unterschiedliche Verhalten in Situationen desselben Typs verantwortlich sind. In diesem Fall wird T_S also Aussagen enthalten wie:

- (2) In Situationen vom Typ A verhält sich S in der Weise X_1 , wenn sich S im Zustand Z_1 befindet.
 In Situationen vom Typ A verhält sich S in der Weise X_2 , wenn sich S im Zustand Z_2 befindet.
 usw.

Die Annahme, daß S verschiedene theoretisch-funktionale Zustände annehmen kann, wird jedoch erst dann systematisch fruchtbar, wenn man über diese Zustände mehr sagen kann, als daß sie das unterschiedliche Verhalten von S in Situationen desselben Typs erklären. Wirklich interessant wird eine entsprechende Verhaltenstheorie, wenn sie (a) auch Aussagen darüber enthält, unter welchen Bedingungen S welche theoretisch-funktionalen Zustände einnimmt und wie diese Zustände untereinander interagieren, und wenn sie (b) zur Erklärung des Verhaltens von S mit einem möglichst kleinen Set von verschiedenen Zuständen auskommt, weil in ihr die Unterschiede im Verhalten von S eher auf die Interaktion schon bekannter als auf immer neue theoretisch-funktionale Zustände zurückgeführt werden. Die Aussagen einer befriedigenden Verhaltenstheorie werden daher in der Regel komplexer sein als Aussagen der Art (1) und (2). Eine solche Verhaltenstheorie wird z.B. auch Aussagen der folgenden Art enthalten:

- (3) Wenn S sich im Zustand Z_1 befindet und in eine Situation vom Typ A kommt, wird S vom Zustand Z_1 in die Zustände Z_2 und Z_3 wechseln.
 (4) Wenn S sich in den Zuständen Z_1 und Z_2 befindet, wird sich S in der Weise X_1 verhalten und vom Zustand Z_2 in den Zustand Z_3 übergehen.

Die Details müssen uns hier jedoch nicht interessieren. Wichtig ist nur, daß auch externe Verhaltenstheorien in der Regel nicht ohne die Annahme auskommen, daß das System, dessen Verhalten erklärt werden soll, verschiedene theoretisch-funktionale Zustände annehmen kann.

Dieser Punkt ist deshalb wichtig, weil er zu der für die Argumentation Dennetts zentralen Frage führt, was man eigentlich erwarten darf, wenn man, nachdem man zunächst eine externe Theorie zur Erklärung des Verhaltens von S entwickelt hat, nachträglich in die Lage versetzt wird, einen *direkten* Einblick in die innere Struktur von S zu erhalten. Kann man erwarten, daß jedem der theoretisch-funktionalen Zustände, die man zur Erklärung des Verhaltens von S postuliert hat, ein struktureller Zustand im inneren Mechanismus von S entspricht?

In manchen Fällen wird das sicher so sein.⁸ Aber theoretisch zwingend ist diese Annahme nicht. Erstens nämlich kann es sein, daß die interne Struktur von S viel komplexer ist als die theoretisch-funktionale Struktur, die man aus dem

Verhalten von *S* erschlossen hatte, und daß den einzelnen theoretisch-funktionalen Zuständen jeweils Klassen von internen Zuständen entsprechen, in denen sehr heterogene interne Zustände zusammengefaßt sind. Zweitens kann es sein, daß sich die interne Struktur von *S* als *funktional undurchdringlich* erweist, so daß es nicht möglich ist, den angenommenen theoretisch-funktionalen Zuständen präzise umschriebene interne Zustände zuzuordnen.⁹ Und drittens schließlich kann es sogar sein, daß man bei der Untersuchung der internen Struktur von *S* zu dem Ergebnis kommt, daß man das Verhalten von *S* unter Bezugnahme auf die internen Zustände von *S* sinnvoll nur erklären kann, wenn man dieses Verhalten mit Hilfe eines Begriffssystems klassifiziert, das sich grundlegend von dem der Theorie *T_S* zugrundeliegenden Begriffssystem unterscheidet.

Mir scheint, daß hinter dem Instrumentalismus Dennetts die Auffassung steht, daß die von einer Verhaltenstheorie *T_S* postulierten theoretisch-funktionalen Zustände genau dann nicht uneingeschränkt real sind, wenn einer dieser Fälle vorliegt, d.h. wenn das System dieser Zustände nicht zu den internen Strukturen paßt, die sich bei der Analyse des 'Innenlebens' von *S* ergeben. Bei komplexen Systemen reicht deshalb der explanatorische und prognostische Erfolg externer Verhaltenstheorien allein nicht aus, um die Realität der im Rahmen dieser Verhaltenstheorien eingeführten Zustände zu begründen. Bei einer Analyse der internen Mechanismen dieser Systeme muß sich außerdem zeigen, daß es eine genaue Entsprechung zwischen den von außen postulierten theoretisch-funktionalen Zuständen und den internen strukturellen Zuständen des Systems gibt.

Bemerkenswerterweise gibt es in diesem Punkt auch gar keinen Dissens zwischen Dennett und Fodor. Denn, wie ich schon erwähnt hatte, ist auch Fodor der Meinung, daß intentionale Zustände nur dann real sein können (und d.h. für ihn insbesondere, daß intentionale Zustände nur dann die für sie charakteristischen kausalen Kräfte haben können), wenn sie durch korrespondierende physische Zustände realisiert sind. In dieser Auffassung liegt für Fodor sogar einer der entscheidenden Gründe für die Grundthese seiner Repräsentationalen Theorie des Geistes:¹⁰

(RTG) Für jeden Organismus *O* und jeden Typ *A* intentionaler Zustände gibt es eine (funktionale/ computationale) Relation *R*, so daß gilt:
O ist genau dann in einem intentionalen Zustand des Typs *A* mit dem Inhalt *p*, wenn sich *O* in der Relation *R* zu einer mentalen Repräsentation *MR* befindet und *MR* die Bedeutung *p* hat. (Fodor 1987, S. 17)

Damit ist, denke ich, aber auch klar, warum Dennett und Fodor - trotz ihrer Übereinstimmung im Hinblick auf den Erfolg und die Unentbehrlichkeit der intentionalen Alltagspsychologie - im Hinblick auf die Frage der Realität intentionaler Zustände so unterschiedliche Auffassungen vertreten: Fodor ist - und das wird in der These (RTG) überdeutlich - davon überzeugt, daß sich z.B. in den Gehirnen von Menschen und Tieren interne Strukturen finden lassen, die ge-

nau den von der intentionalen Alltagspsychologie postulierten intentionalen Zuständen entsprechen. Dennett dagegen hält genau diese Annahme für falsch.

"[...] I am a sort of realist, but I am not Fodor's Realist with a capital R since I expect that the actual internal states that cause behavior will not be functionally individuated, even to an approximation, the way belief/ desire psychology carves things up." (1987d, S. 71)

Die entscheidende Frage ist also, mit welchen Argumenten Fodor und Dennett ihre jeweiligen Positionen zu stützen versuchen.

6. Im Hinblick auf Fodor ist die Beantwortung dieser Frage einfach. Für ihn beruht die These (RTG) einfach auf einem Schluß auf die beste Erklärung. Auch Fodor gesteht zwar zu, daß ein solcher Schluß theoretisch nicht zwingend ist. Aber seiner Meinung nach gibt es einfach keine Alternative, die die Produktivität und Systematizität intentionaler Zustände ebensogut erklären könnte wie die Tatsache, daß die Kausalbeziehungen zwischen intentionalen Zuständen häufig den semantischen Beziehungen entsprechen, die zwischen den Inhalten dieser Zustände bestehen.¹¹

Im Hinblick auf Dennett liegen die Dinge allerdings nicht ganz so einfach. In seinen früheren Aufsätzen gewinnt man häufig den Eindruck, als wolle er die Position vertreten, daß es aus *theoretischen* Gründen unmöglich ist, daß intentionale Zustände durch strukturelle Zustände physischer Systeme realisiert sind. Dennett unterscheidet in diesen Aufsätzen streng zwischen der personalen (intentionalen) und der subpersonalen Ebene der Verhaltensklärung und suggeriert dabei immer wieder, daß es zwischen diesen beiden Ebenen eine grundsätzliche Kluft gibt, die den Übergang von der einen zur anderen prinzipiell unmöglich macht. Es ist allerdings nicht leicht zu verstehen, welche Gründe Dennett für die Annahme dieser unüberbrückbaren Differenz hat.

Einige Punkte lassen sich aber vielleicht doch rekonstruieren. Da ist zum ersten Dennetts These, daß wir bei der Zuschreibung intentionaler Zustände eine Rationalitätsannahme machen müssen, von der wir wissen, daß ihr kein physisches System entsprechen kann. Mit anderen Worten: Wenn wir einem System intentionale Zustände zuschreiben, setzen wir auf der einen Seite voraus, daß sich dieses System vollkommen rational verhält; auf der anderen Seite wissen wir aber, daß kein physisches System sich vollkommen rational verhalten kann; also kann es keine Entsprechung zwischen der intentionalen und der subpersonalen Ebene geben.

Ein zweites - offensichtlich verwandtes - Argument versteckt sich meiner Meinung nach in Dennetts These, daß es keine semantischen Maschinen gibt,¹² d.h. daß syntaktische Maschinen (und jedes physische System ist *eo ipso* eine syntaktische Maschine) semantische Maschinen immer nur partiell, aber niemals vollständig imitieren können.

Die Frage ist jedoch, was eigentlich für diese These spricht. Auch das wird in Dennetts Arbeiten nicht ganz klar. Mir scheint jedoch, daß Dennett davon ausgeht, daß wir immer, wenn wir eine Maschine als eine semantische Maschine charakterisieren, dieser Maschine eine Fähigkeit oder Kompetenz zuschreiben,

von der aus prinzipiellen Gründen ausgeschlossen ist, daß ein physisches System über diese Fähigkeit bzw. Kompetenz uneingeschränkt verfügt.

Dies zeigt sich z.B. an der Art, in der Dennett in (1987d) und (1987e) das Beispiel eines Taschenrechners verwendet. Bevor ich das genauer erläutern kann, muß ich zuvor jedoch kurz auf die Marrsche Unterscheidung zwischen drei Ebenen der Analyse komplexer kognitiver Systeme¹³ eingehen, da sich Dennett in diesem Zusammenhang häufig auf diese Unterscheidung bezieht. Die höchste der von Marr unterschiedenen Ebenen ist die computationale Ebene. Auf dieser Ebene geht es darum festzustellen, welche input-output-Funktion ein System realisiert, d.h. welchen output es bei welchem input liefert.¹⁴ (Wenn es sich um ein symbolverarbeitendes System handelt, läßt sich diese input-output-Funktion als eine Funktion im arithmetischen Sinne beschreiben. Dies ist wohl der Grund dafür, daß Marr diese Ebene die Ebene der computationalen Analyse nennt.) Auf der zweiten, der algorithmischen Ebene geht es um die Frage, wie die auf der computationalen Ebene gefundene input-output-Funktion berechnet werden kann, d.h. erstens, welches die geeigneten Repräsentationen für den input und output sind, und zweitens, welcher Algorithmus die erforderliche Umformung leistet. Auf der dritten, der Implementationsebene schließlich geht es um die Frage, wie dieser Algorithmus physikalisch realisiert ist.

Für Dennetts Beispiel ist entscheidend, daß wir das Verhalten eines Taschenrechners auf der computationalen Ebene mit Hilfe eines *Kompetenzmodells* analysieren. Denn auf dieser Ebene unterstellen wir schlicht, daß der Rechner im Hinblick auf einige einfache Funktionen wie die Addition, Subtraktion, Multiplikation usw. die Fähigkeit besitzt, für alle Argumente die *richtigen* Funktionswerte zu errechnen. D.h., auf der computationalen Ebene beschreiben wir das Verhalten des Rechners einfach mit Hilfe der Gesetze der Arithmetik. Tatsächlich kann jedoch kein physisches System diese Gesetze uneingeschränkt realisieren. Denn in der Arithmetik spielt Endlichkeit keine Rolle; physische Systeme sind aber, was immer sie sonst sein mögen, auf jeden Fall endlich. Und dies führt dazu, daß kein physisches System für alle möglichen Argumente das jeweils richtige Ergebnis angeben kann.

Dies wird schon bei einfachen Divisionsaufgaben deutlich, bei denen der Rechner nur Näherungswerte angibt, da er entweder die Berechnung bei einer bestimmten Stellenzahl hinter dem Komma abbricht oder aufrundet. Im ersten Fall erhält man bei der Division von 2 durch 3 das 'Ergebnis' 0,666666, im zweiten Fall das 'Ergebnis' 0,666667. Außerdem sagt uns die Arithmetik, daß 10 geteilt durch 3 multipliziert mit 3 wieder 10 und daß 20 geteilt durch 3 multipliziert mit 3 wieder 20 ergibt, aber viele Taschenrechner liefern im ersten Fall das 'Ergebnis' 9,999999 und im zweiten Fall das 'Ergebnis' 19,999999.

Natürlich können wir auf der Grundlage solcher Beobachtungen versuchen, durch schrittweise Anpassung des reinen Kompetenzmodells zu einem präziseren *Performanzmodell* zu gelangen. Aber dies wäre Dennett zufolge nicht unbedingt vernünftig. Denn erstens müssen wir bei dem Versuch, ein Performanzmodell zu erarbeiten, von der computationalen auf die algorithmische Ebene herabsteigen,¹⁵ wobei wir immer zugleich ein paar der uns von der computationalen Ebene her geläufigen Kategorien verlieren.¹⁶ Und zweitens spre-

chen häufig auch praktische Gründe dafür, am Kompetenzmodell festzuhalten. Wenn nämlich auch die Anwendung des idealisierten Kompetenzmodells einigermaßen verlässliche Verhaltensprognosen liefert und wenn die Anwendung dieses Modells zugleich wesentlich einfacher ist als die des Performanzmodells, dann kann es durchaus sinnvoll sein, das Kompetenzmodell dem Performanzmodell vorzuziehen.

"Moving away from an oversimple ideal toward greater realism is not always a wise tactic. It depends on what you want; sometimes quick-and-dirty prediction is more valuable than an extension of one's fine-grained scientific understanding even in science." (1987d, S. 79)

Allerdings sollte einem dabei Dennett zufolge immer klar sein, daß sich aus einem reinen Kompetenzmodell in der Regel eben keine vernünftigen Vermutungen über die internen Strukturen des betreffenden Systems ableiten lassen.

"The fact about competence models that provokes my 'instrumentalism' is that the decomposition of one's competence model into parts, phases, states, steps, or whatever *need* shed no light at all on the decomposition of actual mechanical parts, phases, states, or steps of the system being modeled - even when the competence model is, as an competence model, excellent." (1987d, S. 75 f.)

Diese Argumentation hat jedoch eine Reihe von schwachen Punkten. Zunächst einmal ist Dennetts Beispiel offensichtlich nicht optimal gewählt. Denn gerade bei Taschenrechnern ist das Kompetenzmodell ein recht guter Ausgangspunkt für die Analyse der internen Strukturen. Die Arbeitsweise solcher Rechner beruht nämlich tatsächlich auf Algorithmen, die im Prinzip zu den arithmetisch richtigen Resultaten führen. Falsche oder nur annähernd richtige Ergebnisse beruhen in diesem Zusammenhang allein darauf, daß wegen der endlichen Größe der Register des Rechners diese Algorithmen in vielen Fällen vorzeitig abgebrochen werden müssen.

Aber lassen wir diesen Punkt beiseite. Der zweite unklare Punkt besteht meiner Meinung nach darin, daß Dennett unterstellt, daß wir bei dem Versuch, aus einem idealisierten Kompetenzmodell ein realistischeres Performanzmodell zu entwickeln, immer von der computationalen auf die algorithmische Ebene herabsteigen müssen und daß wir aus diesem Grund bei einem solchen Versuch zugleich immer ein paar der uns von der computationalen Ebene her geläufigen Kategorien verlieren. Beides scheint mir nicht zwingend. Bei dem Versuch, für einen Taschenrechner statt eines Kompetenz- ein Performanzmodell zu entwickeln, können wir ohne weiteres auf der computationalen Ebene bleiben, indem wir einfach ein Stück 'Taschenrechnerarithmetik' entwickeln: eine Arithmetik, derzufolge z.B. das Ergebnis von 10 geteilt durch 3 nicht $3\frac{1}{3}$, sondern 3,333333 ist. Und selbst wenn wir von der computationalen auf die algorithmische Ebene hinabsteigen, bedeutet das keineswegs immer, daß wir dabei das von der computationalen Ebene her vertraute Begriffssystem verändern müssen. Die Funktion der Multiplikation von x und y z.B. kann mit Hilfe des folgenden Algorithmus durchgeführt werden:

1. Setze das Ergebnisregister z auf Null.
2. Wenn y gleich Null ist, beende den Prozeß.
3. Addiere x zu z und vermindere y um 1. Gehe zu 2.

Und zur Beschreibung dieses Algorithmus verwenden wir offenbar das gleich arithmetische Begriffssystem, mit dem wir auch die zu berechnende Funktion beschrieben hatten.

Wenn diese Einwände zutreffen, scheint mir aber nichts dagegen zu sprechen, daß wir uns - zumindest in vielen Fällen - bei dem Versuch, eine externe Theorie zur Erklärung des Verhaltens eines Systems zu entwickeln, statt an Kompetenzmodellen an realistischeren Performanzmodellen orientieren können, ohne deshalb die Beschreibungsebene wechseln zu müssen. Performanzmodelle können aber physisch realisiert sein. Also gibt es in diesen Fällen zumindest keinen a priori Grund mehr für die Annahme, daß die internen Strukturen der jeweiligen Systeme nicht den in der Verhaltenstheorie unterstellten theoretisch-funktionalen Strukturen entsprechen.

Dennett könnte jedoch darauf bestehen, daß wir es wenigstens im Falle intentionaler Erklärungen mit Erklärungen zu tun haben, die auf einem Kompetenz- und nicht auf einem Performanzmodell beruhen.¹⁷ Aber auch diese Auffassung ist zweifelhaft. Denn sie beruht letzten Endes auf der schon für das erste Argument zentralen Annahme, daß wir intentionale Zustände nur unter der Voraussetzung völliger Rationalität zuschreiben können. Und diese Annahme ist selbst eher unplausibel. Dennett hat zwar des öfteren die Auffassung vertreten, die intentionale Psychologie sei eigentlich eine normative Theorie, die uns sage, was jemand glauben und welche Wünsche er haben solle und was für jemanden, der diese Überzeugungen und Wünsche habe, zu tun rational sei.¹⁸ Aber die intentionale Alltagspsychologie, so wie wir sie kennen, und auch die intentionale kognitive Psychologie verstehen sich ganz im Gegenteil als normale empirische Theorien, die uns sagen, welche Wünsche und Überzeugungen Personen tatsächlich haben und inwieweit ihr Verhalten in Bezug auf diese Wünsche tatsächlich rational bzw. irrational ist. Natürlich können auch die intentionale Alltagspsychologie und die kognitive Psychologie nicht davon ausgehen, daß sich Personen völlig irrational verhalten. Denn in diesem Fall wäre der Zuschreibung intentionaler Zustände in der Tat jede Grundlage entzogen. Aber das Gegenteil von 'völlig irrational' ist nicht 'völlig rational'. Und tatsächlich gehört es zu den interessantesten Aufgaben der intentionalen Psychologie zu erkunden, wieweit der Bereich der Rationalität reicht und unter welchen Bedingungen Personen dazu neigen, falsche Überzeugungen zu erwerben, falsche Schlüsse zu ziehen oder irrational zu handeln. D.h., sowohl die intentionale Alltagspsychologie als auch die kognitive Psychologie verstehen sich zu Recht nicht als Kompetenz-, sondern als Performanztheorien. Und offensichtlich können sie das tun, ohne deshalb das intentionale Begriffssystem verlassen zu müssen.¹⁹

7. Die a priori Argumente Dennetts sind also nicht sehr überzeugend. Es ist allerdings nicht auszuschließen, daß Dennett seine Position in diesem Punkt inzwischen geändert hat. Denn in seinen letzten Aufsätzen scheint er an mehreren Stellen die Auffassung zu vertreten, daß es sich hier eigentlich um eine empirische Frage handelt und nicht um eine Frage, die sich mit Hilfe von a priori Argumenten entscheiden läßt. Bemerkenswert ist z.B., daß Dennett in (1987d) die folgende Bemerkung von Bill Bechtel offenbar zustimmend zitiert:²⁰

"[...] Dennett seems to reduce the question of instrumentalism versus realism to an *empirical* issue about how the human cognitive system is structured - if there turns out to be a reasonable mapping of intentional idioms onto processing states, then realism will be vindicated, while instrumentalism is justified if there is no such mapping." (1985, S. 477; Hervorh. vom Verf.)

Noch deutlicher wird dieser Sinneswandel jedoch in den in (1987e) vorgetragenen Bemerkungen zur Geschichte und zu den Erfolgsaussichten der Fodorschen Theorie einer Sprache des Geistes. Dennett beginnt diese Bemerkungen mit der Feststellung, daß Fodors Theorie von Anfang an eher durch apriorische Überlegungen zur Frage, welche Strukturen den kognitiven Leistungen von Tieren und Menschen überhaupt zugrundeliegen *können*, als durch präzise empirische Befunde gestützt wurde. Und dann fährt er fort (wieder bemerkenswerterweise, wie mir scheint):

"As more light has been shed on the mechanisms and methods of various psychological and neural systems, it is emerging neither vindicated nor entirely discredited. It would have been vindicated with fanfare had it led to detailed and testable hypotheses about the organization of cognitive processes, but nothing like that has happened." (S. 227)

Um den gegenwärtigen Status der Theorie einer Sprache des Geistes angemessen beurteilen zu können, sei es, so Dennett, hilfreich, einmal zu fragen, unter welchen Bedingungen diese Theorie eigentlich als eine erfolgreiche Theorie gelten würde. Bei seiner Formulierung dieser Bedingungen bezieht sich Dennett wieder auf die Marrsche Unterscheidung zwischen der *computationalen Ebene*, der *algorithmischen Ebene* und der *Implementationsebene*. Vor dem Hintergrund dieser Unterscheidung formuliert er drei Bedingungen, die seiner Meinung nach erfüllt sein müßten, um Fodors Theorie zu einer erfolgreichen Theorie zu machen:

1. Die Prinzipien der *intentionalen Alltagspsychologie* müßten sich so präzise formulieren lassen wie die Prinzipien der *Arithmetik*. Und es müßte sich zeigen lassen, daß diese Prinzipien das Verhalten von Menschen und höheren Säugetieren ebenso gut beschreiben, wie die Gesetze der *Arithmetik* das Verhalten von Taschenrechnern. D.h. die Prinzipien der *intentionalen Alltagspsychologie* müßten eine *adäquate Beschreibung* auf der *computationalen Ebene* liefern.

2. Es müßte sich zeigen, daß es beim Abstieg von der *computationalen* auf die *algorithmische Ebene* nicht nötig ist, das *Begriffssystem* radikal zu wechseln. D.h. die Zustände, die die Ausgangs- und Endpunkte der angenommenen Algorithmen bilden, müßten zumindest eine *starke Ähnlichkeit* zu den auf der *computationalen Ebene* angenommenen Zuständen haben.

3. Die angenommenen Algorithmen müßten sich in *natürlicher Weise* auf die *neuronale hardware* abbilden lassen. D.h., die im Gehirn stattfindenden neuronalen Prozesse müßten sich *zwanglos* als *Implementierungen* dieser Algorithmen interpretieren lassen.²¹

Diese Bedingungen entsprechen offenbar genau den Überlegungen, auf die ich schon im letzten Abschnitt eingegangen bin. Der entscheidende Punkt ist meiner Meinung nach aber, daß Dennett an dieser Stelle nicht mehr argumentiert, daß die genannten drei Bedingungen aus *apriorischen Gründen* nicht erfüllt

sein können. Stattdessen verweist er einfach darauf, daß *de facto* nichts dafür spricht, daß sie erfüllt sind. Das zeigt sich Dennett zufolge an zwei Punkten besonders deutlich. Erstens daran, daß der Fortschritt in den Gebieten, in denen die Kognitionswissenschaft bisher tatsächlich erfolgreich war, (nämlich im Bereich der peripherienahen sensorischen und motorischen Systeme) vielfach gerade darin bestand, daß gezeigt werden konnte, wie solche Systeme ihre Aufgabe erfüllen können, ohne daß Symbolverarbeitungsprozesse im traditionellen Sinne dabei irgendeine Rolle spielen. Und zweitens daran, daß es im Bereich der zentralen kognitiven Aktivitäten, von dem man am ehesten denken würde, daß es in ihm ohne Symbolverarbeitungsprozesse nicht geht, in der Zwischenzeit kaum theoretische und empirische Fortschritte gegeben hat.

"The problem is not that sententialist models of such thinking do not lead to plausible testable hypotheses. Worse, they seem to lead quite systematically down recognizable dead ends: hopelessly brittle, inefficient, and unversatile monstrosities of engineering that would scarcely guide an insect through life unscathed." (1987e, S. 229)

Mit einem Wort: Die Idee einer Sprache des Geistes hat sich Dennett zufolge empirisch als ein degeneratives Forschungsprogramm erwiesen. Es bedarf nur noch einer erfolgversprechenden Alternative, um ein weiteres Festhalten an diesem Forschungsprogramm völlig irrational werden zu lassen. Wenn sich der Konnektionismus - wie viele hoffen und einige fürchten - als diese Alternative erweisen sollte, wären die Tage der Theorie einer Sprache des Geistes endgültig gezählt.

8. Wenn meine Interpretation zutrifft, daß Dennett die Frage, ob die intentionale Alltagspsychologie die im letzten Abschnitt angeführten drei Bedingungen erfüllt, inzwischen für eine empirische Frage und nicht mehr für eine Frage hält, über die mit apriorischen Gründen entschieden werden kann, dann sind wir in gewisser Weise offenbar wieder zurück am Ausgangspunkt unserer Argumentation. Denn in diesem Fall könnten wir festhalten, daß alle Beteiligten die Frage nach der Realität intentionaler Zustände für eine empirische Frage halten, eine Frage, über die allein durch den Fortschritt der empirischen Wissenschaften entschieden wird.

Allerdings hat sich die Meinung darüber, welche empirischen Fakten für die Beantwortung dieser Frage entscheidend sind, doch erheblich verändert. Denn so wie sich die Diskussion entwickelt hat, geht es heute nicht mehr in erster Linie darum, ob die intentionale Alltagspsychologie oder die an sie anknüpfende kognitive Psychologie eine erfolgreiche Theorie im Bereich der Erklärung und Voraussage von Verhalten ist. (Das ist zwar immer noch eine notwendige Bedingung; aber es reicht allein nicht mehr aus, um die Realität intentionaler Zustände zu begründen.) Entscheidend ist heute vielmehr die Frage, ob sich empirisch plausibel machen läßt, daß es eine genaue Entsprechung zwischen den von der Alltagspsychologie angenommenen intentionalen Zuständen und den physischen Zuständen in den internen Mechanismen der jeweiligen Systeme gibt. Fodor glaubt, daß das der Fall ist, weil seiner Meinung nach kognitive Leistungen auf andere Weise nicht realisiert sein können. Dennett dagegen ist davon über-

zeugt, daß es nicht der Fall ist, und er vertritt diese Position insbesondere deshalb, weil sich seiner Meinung nach die Idee einer Sprache des Geistes empirisch als ein degeneratives Forschungsprogramm erwiesen hat. Dies scheint mir, ist der entscheidende Grund für seine Ablehnung der realistischen Deutung intentionaler Zustände.²²

Bleibt zum Schluß allerdings noch die Frage, warum Dennett sich dann nicht auf die Seite der Churchlands schlägt, die aus ziemlich denselben empirischen Befunden den Schluß ziehen, daß es intentionale Zustände einfach nicht gibt und daß nur die internen strukturellen Zustände real sind. Die Antwort auf diese Frage ist ganz offensichtlich in Dennetts Überlegungen zum Thema 'real patterns' zu suchen. Aber leider gibt es in diesen Überlegungen eine merkwürdige Zweideutigkeit. Bei den Beispielen, die Dennett in diesem Zusammenhang anführt und auf die ich in den Abschnitten 2 und 4 eingegangen bin, handelt es sich eigentlich immer um Muster im output, d.h. im Verhalten eines Systems. An einigen Stellen spricht Dennett jedoch auch von Mustern von internen Zuständen, also z.B. von Mustern von Wünschen und Überzeugungen.²³ Ich denke allerdings, daß Dennetts Argumentation nur dann einen vernünftigen Sinn hat, wenn sie im Sinne der ersten Möglichkeit verstanden wird. Denn Dennetts These, daß die Besucher vom Mars bestimmte reale Muster nicht wahrnehmen können, wenn sie das Verhalten von Menschen und Tieren nicht in intentionaler Einstellung erklären, ist trivial und nichtssagend, wenn es sich bei diesen Mustern um Muster von intentionalen Zuständen handeln soll. Die Muster, um die es Dennett geht, versteht man deshalb meiner Meinung nach am besten als Verhaltensregelmäßigkeiten, die sich erst zeigen, wenn man das Verhalten mit Hilfe eines bestimmten Begriffssystems beschreibt.

Wenn das so ist, läßt sich die Argumentation Dennetts gegen den Eliminativismus der Churchlands aber folgendermaßen rekonstruieren. Auf der einen Seite stimmt Dennett mit den Churchlands darin überein, daß sich auf der neuronalen Ebene keine Entsprechungen zu den von der Alltagspsychologie postulierten intentionalen Zuständen finden lassen. Auf der anderen Seite steht er der Churchlandschen Auffassung jedoch sehr skeptisch gegenüber, daß die neuen exakteren neurobiologischen Verhaltensklärungen sehr schnell die herkömmlichen unpräzisen intentionalen Erklärungen ersetzen werden, wenn wir erst einmal richtig verstanden haben, welche neurobiologischen Strukturen dem Verhalten von Menschen und Tieren zugrundeliegen. Warum? Offenbar ist Dennett der Meinung, daß diese neurobiologischen Strukturen nicht zu dem Begriffssystem passen, mit dem wir Verhalten klassifizieren, wenn wir es intentional erklären. Oder mit anderen Worten: Dennett ist davon überzeugt, daß sich Verhalten *als Handeln* nicht neurobiologisch erklären läßt. Wenn wir von der Ebene intentionaler Zustände auf die Ebene neuronaler Strukturen herabsteigen, verlieren wir die Möglichkeit zu erklären, warum Hans Paul ermahnt, vorsichtig zu sein, oder warum er jetzt in sein Büro geht. Auf der neuronalen Ebene können keine Handlungen, sondern nur Körperbewegungen erklärt werden. Auf dieser Ebene können wir daher bestenfalls erklären, warum sich Hans' Mund und Stimmbänder soundso bewegen oder warum jetzt die Muskeln kontrahieren, die seine Beine in Bewegung setzen.

Ich bin mir nicht hundertprozentig sicher, daß dies die korrekte Rekonstruktion der Argumentation Dennetts ist. Besonders die letzten beiden Seiten von (1991) scheinen mir in diesem Punkt nicht völlig klar zu sein. Trotzdem denke ich, daß diese Interpretation gut in den Gesamtrahmen der Dennettschen Überlegungen paßt. Aber sie führt natürlich zu zwei weiteren Fragen. Die erste Frage ist, welche Gründe denn für die These sprechen, daß auf der neuronalen Ebene keine Handlungen, sondern nur Körperbewegungen erklärt werden können. Und die zweite Frage lautet, warum die Klassifikation von Körperbewegungen als Handlungen für Dennett einen solchen Stellenwert hat, daß er sie für praktisch unauflösbar hält.

Im Hinblick auf die zweite Frage ist klar, daß Dennett wieder auf seine 'real patterns' Überlegungen verweisen würde. Wenn wir Körperbewegungen als Handlungen klassifizieren, erfassen wir real existierende Muster, die uns verborgen bleiben, wenn wir nur die "bit map" der einzelnen Körperbewegungen betrachten. Und wenn wir Körperbewegungen mit Hilfe dieses Musters klassifizieren, erhalten wir darüber hinaus einen enormen praktischen Vorteil, da wir in diesem Fall zu viel effizienteren, wenn auch nicht ganz so präzisen Vorhersagen kommen. In dieser Argumentation klafft jedoch eine Lücke. Denn sie unterstellt, daß es kein alternatives Muster gibt, das auf der einen Seite genauso effizient ist wie das Handlungsmuster, auf der anderen Seite aber besser zu den neuronalen Strukturen in unseren Gehirnen paßt. Denn wenn es ein solches Muster gäbe, könnten wir in ebenso effizienter Weise (nach diesem Muster klassifizierte) Körperbewegungen durch die Bezugnahme auf neuronale Strukturen erklären und voraussagen, ohne (zusätzliche) Zustände postulieren zu müssen, die nicht zu diesen Strukturen passen.

Wie steht es schließlich mit der ersten These, daß auf der neuronalen Ebene keine Handlungen, sondern nur Körperbewegungen erklärt werden können? Nun, man könnte sagen, daß diese These zumindest eine große *prima facie* Plausibilität hat. Und man könnte auch sagen, daß der Begriff der Handlung den Begriff des intentionalen Zustands schon voraussetzt, da nur mit Bezugnahme auf entsprechende intentionale Zustände eine Körperbewegung überhaupt als Handlung bzw. als Handlung eines bestimmten Typs klassifiziert werden kann. Doch mit dieser Antwort würde man möglicherweise nur den Verdacht nähren, daß Handlungen eben doch kein reales Muster, sondern nur ein Artefakt darstellen, das nur für den sichtbar wird, der sich zugleich auf die in ihrer Realität unstrittenen intentionalen Zustände einläßt. Doch davon einmal abgesehen. Mir scheint auch, daß diese These dem Selbstverständnis der Neurobiologie nicht gerecht wird. Denn die Neurobiologie interessiert sich keineswegs nur für die Innervation einzelner Muskeln. Ihr geht es vielmehr gerade auch um die Frage, wie es das Gehirn und das ZNS zustandebringen, daß wir aufrecht stehen und laufen können bzw. welche neuronalen Strukturen und Prozesse komplexen koordinierten Bewegungen wie dem Greifen oder dem Jagen zugrundeliegen. D.h., der Neurobiologie geht es auch um die Erklärung von komplexen Bewegungsabläufen, die zumindest der philosophisch Unverbildete ohne weiteres als Handlungen bezeichnen würde. Ich glaube nicht, daß man der Neurobiologie mit apriorischen

Gründen glaubhaft machen kann, daß dieses Unterfangen von vornherein zum Scheitern verurteilt ist.

Literatur

- Bechtel, W., 1985, Realism, Instrumentalism, and the Intentional Stance. In: *Cognitive Science* 9, S. 473-497.
- Beckermann, A., 1986, Dennetts Stellung zum Funktionalismus. In: *Erkenntnis* 24, S. 309-341.
- Beckermann, A., 1991, Der endgültige Todesstoß für den Repräsentationalismus? Eine Replik auf Andreas Kemmerlings Artikel 'Mentale Repräsentationen'. In: *Kognitionswissenschaft* 2, S. 91-98.
- Churchland, P.M., 1981, Eliminative Materialism and the Propositional Attitudes. In: *Journal of Philosophy* 78, S. 67-90. Wiederabgedr. in: Churchland, P.M., 1989, *A Neurocomputational Perspective*. Cambridge, Mass.: MIT Press, S. 1-22.
- Davidson, D., 1989, What is Present to the Mind?. In: *Grazer Philosophische Studien* 36, S. 3-18.
- Dennett, D., 1971, Intentional Systems. In: *Journal of Philosophy* 68, S. 87-106. Wiederabgedr. in: Dennett 1978, S. 90-108.
- Dennett, D., 1978, *Brainstorms*. Montgomery, Verm.: Harvester Press.
- Dennett, D., 1981a, True Believers. The Intentional Strategy and Why It Works, in: Heath, A.F. (ed.), *Scientific Explanation*. Oxford. Wiederabgedr. in: Dennett 1987a, S. 13-35.
- Dennett, D., 1981b, Three Kinds of Intentional Psychology, in: Healy, R. (ed.), *Reduction, Time, and Reality*. Cambridge. Wiederabgedr. in: Dennett 1987a, S. 43-68.
- Dennett, D., 1987a, *The Intentional Stance*. Cambridge, Mass.: MIT Press.
- Dennett, D., 1987b, Setting off on the Right Foot, in: Dennett 1987a, S. 1-11.
- Dennett, D., 1987c, Real Patterns, Deeper Facts, and Empty Questions, in: Dennett 1987a, S. 37-42.
- Dennett, D., 1987d, Instrumentalism Reconsidered, in: Dennett 1987a, S. 69-81.
- Dennett, D., 1987e, The Language of Thought Reconsidered, in: Dennett 1987a, S. 227-235.
- Dennett, D., 1987f, Mid-Term Examination: Compare and Contrast, in: Dennett 1987a, S. 339-350.
- Dennett, D., 1991, Real Patterns. In: *Journal of Philosophy* 88, S. 27-51.
- Fodor, J.A., 1979, Three Cheers for Propositional Attitudes, in: Cooper/ Walker (eds.), *Sentence Processing*. Lawrence Erlbaum Associates. Revidierte und erweiterte Fassung in Fodor 1981, S. 100-123.
- Fodor, J.A., 1981, *Representations*. Cambridge, Mass.: MIT Press.
- Fodor, J.A., 1987, *Psychosemantics*. Cambridge, Mass.: MIT Press.
- Marr, D., 1982, *Vision*. New York: W.H. Freeman and Company.
- Parfit, D., 1984, *Reasons and Persons*. Oxford: Oxford University Press.

- Putnam, H., 1983, Computational Psychology and Interpretation Theory, in: Putnam, H. Realism and Reason. Philosophical Papers. Vol. 3. Cambridge: Cambridge University Press, S. 139-154.
- Quine, W.V.O., 1951, Two Dogmas of Empiricism. Philosophical Review. Wiederabgedr. in: Quine, W.V.O. 1961, From a Logical Point of View, 2. Aufl., New York: Harper and Row, S. 20-46.
- Quine, W.V.O., 1960, Word and Object. Cambridge, Mass.: MIT Press.
- Quine, W.V.O., 1970, On the Reasons for Indeterminacy of Translation. In: Journal of Philosophy 77, S. 178-183.
- Ramsey, W., Stich, S. und Garon, J., 1990, Connectionism, Eliminativism, and the Future of Folk Psychology, in: Tomberlin, J. (ed.), Philosophical Perspectives 4. Action Theory and Philosophy of Mind. Atascadero, Cal.: Ridgeview, S. 499-533.

Anmerkungen

- 1 Vgl. besonders 1981.
- 2 Vgl. etwa die folgende Passage aus 1991:
 "We use folk psychology - interpretation of each other as believers, wanters, intenders, and the like - to predict what people will do next. Prediction isn't the only thing we care about, of course. Folk psychology helps us understand and empathize with others, organize our memories, interpret our emotions, and flavor our vision in a thousand ways, but at the heart of all these is the enormous predictive leverage of folk psychology. Without its predictive power, we could have no interpersonal projects or relations at all; human activity would be just so much Brownian motion; we would be baffling ciphers to each other and to ourselves - we could not even conceptualize our own failings. In what follows I will concentrate always on folk-psychological prediction, not because I make the mistake of ignoring all the other interests we have in people aside from making bets on what they will do next, but because I claim that our power to *interpret* the actions of others depends on our power - seldom explicitly exercised - to predict them." (S. 29)
- 3 Vgl. jedoch Bechtel 1985, S. 479.
- 4 Vgl. 1987f, S. 348 ff. und 1991, S. 46 ff.
- 5 In 1987c, S. 41 verweist Dennett auf Parfit, der in 1984 ebenfalls die Quinesche Argumentationsstrategie benutzt, um zu zeigen, daß es auch im Hinblick auf das Problem der personalen Identität Fragen gibt, deren Beantwortung nicht durch die Tatsachen determiniert ist.
- 6 Diese Position ist von Fodor in 1979, S. 116 ff. in der Tat explizit vertreten worden.
- 7 Zu dieser Unterscheidung vgl. auch Beckermann 1986.
- 8 Das elektronische System, das ich in 1986, S. 319 ff. diskutiere, ist offensichtlich ein Beispiel für diesen Fall.
- 9 Diese Möglichkeit ist besonders dann gegeben, wenn sich *S* zur gleichen Zeit in verschiedenen funktionalen Zuständen befinden kann. (So wie wir ja zur gleichen Zeit sehr verschiedene Überzeugungen, Wünsche und noch andere intentionale Zustände haben können.) In diesem Fall ist es nämlich durchaus denkbar, daß die beiden internen Gesamtzustände, in denen sich *S* z.B. befindet, wenn es gleichzeitig in den funk-

tionalen Zustände Z_j und Z_k bzw. gleichzeitig in den funktionalen Zustände Z_j und Z_k ist, keinen gemeinsamen Teilzustand umfassen, der dem einzelnen funktionalen Z_j entspräche. Dieser Punkt ist insbesondere von konnektionistischer Seite immer wieder betont worden.

In 1990 haben z.B. Ramsey, Stich und Garon ein Netzwerk vorgestellt, von dem man mit einigem Recht behaupten kann, daß in ihm "Wissen" über die Wahrheit bzw. Falschheit von 16 Aussagen gespeichert ist. Es gibt jedoch keine Teile in diesem Netzwerk, die man eindeutig den einzelnen Aussagen zuordnen könnte. Dies wird besonders deutlich, wenn man dieses Netzwerk so trainiert, daß es nicht nur auf die ursprünglichen 16, sondern auch auf eine zusätzliche siebzehnte Aussage mit der korrekten Antwort reagiert. Denn bei diesem Training entwickeln sich die Gewichte der Verbindungen in diesem Netzwerk so, daß sie sich am Ende drastisch von den Gewichten der Verbindungen im ursprünglichen Netzwerk unterscheiden. Obwohl das zweite Netzwerk also die Informationen, die in dem ersten Netzwerk enthalten waren, als Teilmenge enthält, gibt es keinen gemeinsamen Teilzustand der beiden Netzwerke, der genau dieser Teilmenge entspräche.

10 Vgl. hierzu Beckermann 1991.

11 Eine ausführliche Rekonstruktion dieser Argumentation findet sich in Beckermann 1991.

12 Besonders in 1981b.

13 Vgl. Marr 1982, S. 24 ff.

14 Für Marr sind Untersuchungen auf der computationalen Ebene häufig normativ, d.h. für ihn geht es bei diesen Untersuchungen weniger um die Frage, welcher output bei gegebenem input erzeugt wird, als vielmehr um die Frage, welcher output erzeugt werden muß, wenn ein gegebenes Ziel erreicht werden soll. Darin spiegelt sich jedoch nur Marrs ingenieurwissenschaftliche Ausgangsposition.

15 "Here I can use the initial competence model to generate hypotheses the falsification of which sheds light on the actual organization of the machinery. This permits a refinement of the competence model itself - in direction of turning it into a performance model. To ask which of the many possible algorithms for approximate arithmetic is used by the calculator is to descend to the algorithmic level." (Vgl. 1987d, S. 78)

Dennetts Darstellung des Taschenrechnerbeispiels ist hier außerordentlich verwirrend. *De facto* beschreibt er nämlich schon das Kompetenzmodell eines solchen Rechners auf der algorithmischen Ebene; er glaubt dabei aber offenbar, sich mit seiner Beschreibung immer noch auf der computationalen Ebene zu befinden.

16 "[...] when we descend to the algorithmic level, we lose a few of the categories of the computational level [...]" (1987d, S. 78).

17 In der Tat spielt diese These in Dennetts Argumentation eine große Rolle. Vgl. z.B. 1981b, S. 58 ff.

18 Besonders prononciert in 1981b.

19 Auf diesen Punkt hat Fodor schon in 1979 ausdrücklich hingewiesen. Dabei hat er u.a. betont, daß eine Theorie, aus der hervorgeht, unter welchen Bedingungen ein System sich nicht mehr völlig rational verhält, häufig viel interessanter ist als eine, die grundsätzlich von der Rationalität aller Überzeugungen, Wünsche und Handlungen ausgeht. Denn wenn man z.B. versucht, einen Schachcomputer zu schlagen, ist es

natürlich sehr hilfreich zu wissen, welche Art von Drohungen der Computer in der Regel unterschätzt oder übersieht.

In 1987d hat Dennett Fodor vorgeworfen, er mißverstehe den Punkt der Dennettschen Position, wenn er argumentiere, daß es manchmal nützlich sei zu wissen, wann ein System nicht völlig rational sei und ein Schachcomputer (als Führer der schwarzen Steine) z.B. eine Drohung übersähe.

"What is Black not likely to notice? I can tell you without even knowing who Black is. Being approximately (sic!) rational, Black is not likely to notice threats that would take a great deal of time and effort to discover and is extremely likely to notice obvious threats. If Black is, as Fodor supposes, rather unlikely to notice the threat, it must be because the threat is somewhat distant in the search tree and hence may well fall outside Black's more or less optimal focus of attention. So even when we are planning to exploit another rational agent's foibles, we make use of the rationality assumption to guide our efforts." (S. 80)

Mit dieser Erwiderung scheint mir Dennett allerdings Fodor eher noch zusätzliche Argumente zu verschaffen. Denn eigentlich alle Begriffe, die er benutzt (z.B. "taking a great deal of time and effort", "obvious threat", "distant in the search tree", "falling outside the optimal focus of attention"), haben ihren natürlichen Platz in einer Performanz- und nicht in einer Kompetenztheorie. Und viele dieser Begriffe sind darüber hinaus eher auf der Ebene der 'design stance' als auf der Ebene der 'intentional stance' angesiedelt.

- 20 Dies ist umso bemerkenswerter, als Bechtel selbst diese Meinung gar nicht teilt, sondern die Auffassung vertritt, daß Dennetts Instrumentalismus vielmehr auf einer Reihe von nicht-empirischen Argumenten beruht.
- 21 Wenn diese Bedingungen erfüllt wären, spräche Dennett zufolge alles für den intentionalen Realismus:
 "Suppose all this were so; then Fodor and the other Realists would be in heaven, for this would be the reduction of propositional attitude psychology to High Church Computationalism [...]. Beliefs and other propositional attitudes, we would have confirmed, are perfectly real [...]" (1987e, S. 228).
- 22 Meiner Meinung nach müssen auch die Dennettschen Überlegungen, auf die ich im Abschnitt 3 eingegangen bin, in diesem Sinne verstanden werden. Denn daß es Dennett zufolge keine Möglichkeit gibt, zwischen den verschiedenen intentionalen Interpretationen zu entscheiden, die Hans und Paul für das Verhalten Angelikas entwickeln, liegt ganz wesentlich auch daran, daß es seiner Meinung nach eben nicht möglich ist, zwischen diesen Interpretationen dadurch zu entscheiden, daß man gewissermaßen in das Gehirn Angelikas hineinsieht.
- 23 Vgl. etwa 1991, S. 46, 50.