

Der endgültige Todesstoß für den Repräsentationalismus?*

Eine Replik auf Andreas Kemmerlings Artikel „Mentale Repräsentationen“

Ansgar Beckermann

Philosophisches Seminar, Georg-August-Universität, Platz der Göttinger Sieben 5, W-3400 Göttingen, Bundesrepublik Deutschland

The Final Death Blow for Representationalism?

Summary. In his paper *Mentale Repräsentationen*, Andreas Kemmerling claims that representationalism – the assumption that there are mental representations – is bound to fail since even Fodor's version of this theory, the version that is most expressly designed to support this assumption, encounters insurmountable difficulties. In this paper, it is argued that Kemmerling's considerations are unconvincing. In the first part, the main theses of Fodor's account are reconstructed in order to clarify the interconnections between these theses and the arguments that may be advanced to support them. In the second part, it is argued that even if Kemmerling's arguments against Fodor's accounts of broad and narrow content were convincing, this would not imply that representationalism is bound to fail since one may well be an adherent of representationalism without buying any of Fodor's theses concerning the content of mental representations.

Zusammenfassung. In seinem Artikel „Mentale Repräsentationen“ vertritt Andreas Kemmerling die Auffassung, der Repräsentationalismus, d.h. die Annahme, es gebe mentale Repräsentationen, sei zum Scheitern verurteilt. Denn selbst Fodors Version dieser Theorie – die Version, die sich zur Stützung dieser Annahme „am meisten ins Zeug“ werfe – sei mit unüberwindlichen Schwierigkeiten konfrontiert. Meiner Meinung nach sind die Argumente Kemmerlings jedoch nicht überzeugend. Um dies zu zeigen, versuche ich im ersten Teil meiner Replik, die Hauptthesen der Fodorschen Version des Repräsentationalismus zu rekonstruieren, um klarzumachen, welche Zusammenhänge zwischen diesen Thesen bestehen und welche Argumente zu ihrer Stützung angeführt werden können. Im zweiten Teil geht es mir dann darum zu zeigen, daß der Repräsentationalismus selbst dann nicht zum Scheitern verurteilt wäre, wenn Kemmerling mit seinen Argumenten gegen Fodors Thesen zum engen und weiten Inhalt Recht hätte, da man ohne weiteres Anhänger dieser

Theorie sein kann, ohne auch nur eine von Fodors Thesen zum Inhalt mentaler Repräsentationen zu unterschreiben.

In seinem Artikel „Mentale Repräsentationen“, der in Band 1, Heft 2, dieser Zeitschrift erschienen ist, geht es Andreas Kemmerling um die Frage, ob es neben dem unbestrittenen Phänomen der mentalen Repräsentation (im Singular) auch mentale Repräsentationen (im Plural) gibt, d. h. um die Frage, ob es sinnvoll ist, zur Erklärung von intentionalen Zuständen oder zur Erklärung von kognitiven Fähigkeiten mentale Repräsentationen als theoretische Entitäten anzunehmen. Kemmerlings Thema ist also das, was in den letzten Jahren unter dem Namen „Repräsentationale Theorie des Geistes“ häufig und kontrovers diskutiert wurde.

Da Kemmerling davon ausgeht, daß Jerry A. Fodor der herausragende Vertreter des Repräsentationalismus ist, versucht er seine Kritik in einer Auseinandersetzung mit der Fodorschen Version dieser Theorie zu entwickeln. Dabei geht er in zwei Schritten vor. Im ersten Schritt versucht er, eine möglichst klare Darstellung der Grundthesen der Theorie Jerry Fodors zu geben, um dann im zweiten Schritt mit einer Argumentation, die sich besonders auf Fodors Überlegungen zur Bedeutung mentaler Repräsentationen bezieht, zu zeigen, daß diese Theorie zum Scheitern verurteilt ist. Damit will er dann zugleich plausibel gemacht haben, daß „[m]entale Repräsentation . . . ohne mentale Repräsentationen statt[findet]“ (57 a), d. h. daß die Grundidee des Repräsentationalismus verfehlt ist. Ich habe mit beiden Schritten, besonders aber mit dem zweiten, meine Schwierigkeiten und möchte deshalb im folgenden begründen, warum ich nicht glaube, daß Kemmerling mit seinem Angriff auf den Repräsentationalismus tatsächlich Erfolg hat.

1. Meine Einwände gegen den ersten Teil der Argumentation Kemmerlings beziehen sich weniger auf die Thesen, in denen er die Fodorsche Theorie – meiner Meinung nach im wesentlichen korrekt – zusammenfaßt, sondern darauf, daß in Kemmerlings 10-Punkte-Programm nicht deutlich wird, daß diese Thesen keineswegs ein homoge-

* Besonderen Dank schulde ich Peter Lanz für seine sehr hilfreichen Bemerkungen zu einer früheren Fassung dieser Replik.

nes Ganzes bilden, sondern einen sehr unterschiedlichen Status und eine sehr unterschiedliche Bedeutung für die Repräsentationale Theorie des Geistes haben. Besonders bedenklich scheint mir, daß Kemmerling in seiner additiven Aneinanderreihung eine Unterscheidung völlig außer Acht läßt, die für das Verständnis der Fodorschen Position außerordentlich wichtig ist: die Unterscheidung zwischen Thesen, die sich auf die Analyse der Alltagspsychologie und der von dieser Theorie postulierten intentionalen Zustände beziehen, und Thesen, in denen es sozusagen um die Implementationsebene geht, d. h. darum, wie Systeme beschaffen sein müssen, deren Verhalten sich realistisch in intentional alltagspsychologischer Weise erklären läßt. Dies ist deshalb besonders mißlich, weil damit auch die Gründe im Dunkeln bleiben, die Fodor zufolge für die Repräsentationale Theorie des Geistes sprechen.

Die Hauptthese der Fodorschen Analyse der intentionalen Alltagspsychologie lautet:

(RT) Die intentionale Alltagspsychologie ist eine außerordentlich erfolgreiche, tiefgehende und letztlich unaufgebbare Theorie. Wir müssen daher davon ausgehen, daß die von dieser Theorie postulierten intentionalen Zustände wirklich existieren und daß sie tatsächlich die kausale Wirksamkeit haben, die ihnen in den Gesetzen dieser Theorie zugeschrieben wird (Fodor 1987, 1–16).

So weit, so gut. Aber was folgt aus dieser Realismusthese? Worauf lassen wir uns ein, wenn wir annehmen, daß es intentionale Zustände wirklich gibt? Fodor erläutert die Konsequenzen, indem er die charakteristischen Merkmale anführt, die intentionale Zustände von Zuständen anderer Art unterscheiden. In Thesenform lassen sich diese Charakteristica so formulieren (vgl. bes. Fodor 1978, 1985, 1987):

- (11) Intentionale Zustände sind *prima facie* relational. Sie unterscheiden sich entlang zweier Dimensionen: ihrem Typ und ihrem Inhalt nach.¹
- (12) Wegen ihres Inhaltsaspektes sind intentionale Zustände semantisch bewertbar, d. h. sie haben Wahrheits- bzw. Erfüllungsbedingungen.
- (13) Intentionale Zustände sind opak; d. h. in den daß-Sätzen, die wir zur Bezeichnung ihres Inhalts verwenden, können Teilausdrücke in der Regel nicht *salva veritate* durch extensionsgleiche Ausdrücke ersetzt werden.
- (14) Die Inhalte intentionaler Zustände haben eine logische Form.²
- (15) Wie man an den Gesetzen der intentionalen Psychologie ersehen kann, entsprechen den kausalen

¹ Diese Unterscheidung kann man so erläutern: Mein Wunsch, ein neues Fahrrad zu erwerben, und mein Wunsch, einen alten Freund wiederzutreffen, sind intentionale Zustände desselben Typs; beides sind Wünsche. Sie unterscheiden sich jedoch dem Inhalt nach, da das, was gewünscht wird, verschieden ist. Meine Befürchtung, daß es heute regnen wird, und meine Überzeugung, daß es heute regnen wird, sind dagegen intentionale Zustände verschiedenen Typs. Sie haben jedoch denselben Inhalt: die Proposition, daß es heute regnen wird.

² Diese Formulierung aus (1978, 182) kann Fodor eigentlich nicht aufrecht erhalten, weil er, wie Kemmerling (S. 49a) zu Recht betont, die Meinung vertritt, daß Propositionen unstrukturiert sind. Es

Relationen zwischen intentionalen Zuständen häufig semantische Beziehungen zwischen ihren Inhalten.

Für diese These gibt es jedoch eine Alternativformulierung, die das Gemeinte vielleicht noch besser wiedergibt: (15') Den Gesetzen der intentionalen Psychologie und damit den kausalen Relationen zwischen intentionalen Zuständen entsprechen sehr häufig Rationalitätsprinzipien.

Damit ist z. B. gemeint, daß in vielen Fällen gilt: Wenn es rational ist, wenn man *p* und *q* glaubt, auch *r* zu glauben, dann gibt es ein Gesetz der intentionalen Psychologie, das besagt, daß die Überzeugung, daß *p*, und die Überzeugung, daß *q*, zusammen die Überzeugung, daß *r*, verursachen; wenn es rational ist, *X* zu tun, wenn man *p* glaubt und *q* will, dann gibt es ein Gesetz der intentionalen Psychologie, das besagt, daß die Überzeugung, daß *p*, und der Wunsch, daß *q*, die Handlung *X* verursachen; usw.³

Soweit Fodors Analyse der Alltagspsychologie und der von ihr postulierten intentionalen Zustände. Was folgt nun aus dieser Analyse im Hinblick auf die Frage, wie Systeme beschaffen sein müssen, deren Verhalten mit Hilfe von intentionalen Zuständen erklärt werden kann und von denen man daher zu Recht sagen kann, daß sie tatsächlich intentionale Zustände annehmen können? Fodors Antwort auf die Frage scheint mir von der folgenden Grundeinstellung geprägt zu sein:

(PRT) Intentionale Zustände müssen physisch realisiert sein, um kausal wirksam sein zu können.

Denn Fodor zufolge ist es schwer vorstellbar, wie intentionale Zustände z. B. Handlungen (und d. h. auch Körperbewegungen) verursachen können, wenn sie nicht selbst (in welcher Art auch immer) einen physischen Charakter haben.⁴

Im Hinblick auf die Art der physischen Realisierung läßt sich nach Fodor aber noch mehr sagen. Denn aufgrund von (11) scheint es ihm *prima facie* plausibel anzunehmen, daß den beiden Aspekten intentionaler Zustände auch zwei Aspekte ihrer physischen Realisierung entsprechen: dem Inhalt eine mentale Repräsentation und dem Zustandstyp eine bestimmte funktionale/computationale Relation. Dies jedenfalls führt Fodor zu der zentralen These seiner Repräsentationalen Theorie des Geistes:

(RTG) Für jeden Organismus *O* und jeden Typ *A* intentionaler Zustände gibt es eine (funktionale/computationale) Relation *R*, so daß gilt:

scheint mir aber noch unglücklicher, wenn Fodor in (1987, 13) schreibt, intentionale Zustände *selbst* hätten eine Konstituentenstruktur, und wenn Kemmerling dies zur Grundlage der Formulierung seiner These (3) macht (S. 48b). Richtiger wäre es m. E. zu sagen, daß die Gesetze der intentionalen Psychologie häufig unter Bezugnahme auf die logische Struktur der Sätze formuliert werden, mit denen die Inhalte der entsprechenden intentionalen Zustände bezeichnet werden.

³ Offenbar gibt es einen Zusammenhang zwischen den Thesen (14) und (15) bzw. (15'). Denn die Rationalitätsbeziehungen zwischen intentionalen Zuständen lassen sich häufig nur formulieren, wenn man auf die logische Form ihrer Inhalte achtet.

⁴ Fodor ist im Hinblick auf diese Grundüberzeugung wenig explizit. Vgl. aber (1987, 30 und 135).

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, 17).

Mentale Repräsentationen sind dabei als physische Strukturen innerhalb eines Systems zu verstehen, die etwas repräsentieren und insofern eine Bedeutung haben.⁵ Und daß ein System zu einer mentalen Repräsentation in einer bestimmten funktionalen/computationalen Relation steht, bedeutet, daß diese Repräsentation in dem System und insbesondere in den im System ablaufenden Symbolverarbeitungsprozessen⁶ eine bestimmte funktionale Rolle spielt.⁷

Auch in der These (RTG) ist jedoch noch nicht alles enthalten, was man Fodor zufolge über die physische Realisierung intentionaler Zustände sagen kann. Für ihn ist auch die folgende These über die Struktur mentaler Repräsentationen, die eigentliche Language-of-Thought-Hypothese, zumindest prima facie plausibel:

(LT) a. Mentale Repräsentationen sind *strukturiert*; sie haben typischerweise eine Konstituentenstruktur.

⁵ In dieser Hinsicht unterscheiden sich mentale Repräsentationen von den meisten anderen physischen Strukturen. Allerdings nicht von allen: Auch die Äußerungen von Sätzen einer Sprache oder Landkarten oder die Kerben auf dem Griff des Revolvers eines Westernhelden sind physische Strukturen mit Bedeutung. Man kann jedoch nicht davon ausgehen, daß die Tatsache, daß alle diese Strukturen Bedeutungsträger sind, auf einem einzigen gemeinsamen Merkmal beruht. In verschiedenen Fällen können möglicherweise sehr unterschiedliche Umstände dafür verantwortlich sein, daß eine physikalische Struktur *P* eine (bestimmte) Bedeutung hat.

⁶ Kemmerling übersetzt „computational“ häufig mit „algorithmisch“, was mir nicht sonderlich glücklich zu sein scheint. Denn das Entscheidende an der „computational theory of mind“ ist die Annahme, daß die kausalen Beziehungen zwischen intentionalen Zuständen auf Symbolverarbeitungsprozessen beruhen. Ob diese Symbolverarbeitungsprozesse als die Abarbeitung von Algorithmen aufgefaßt werden können, ist dabei völlig zweitrangig.

Auch die Übersetzung von „computational X“ durch „Computer-X“ kann zu Mißverständnissen führen, wenn durch sie die Auffassung hervorgerufen wird, daß Fodor mit seiner Theorie auf die These festgelegt ist, daß die den intentionalen Zusammenhängen zugrunde liegenden Symbolverarbeitungsprozesse nur nach dem Modell der in seriellen von-Neumann-Maschinen ablaufenden Symbolverarbeitungsprozesse verstanden werden können.

⁷ Häufig ist vereinfachend gesagt worden, daß die Relationen, die (RTG) zufolge Wünschen und Überzeugungen entsprechen, einfach darin bestehen, daß sich entsprechende mentale Repräsentationen in der Wunsch- bzw. in der Überzeugungsbox des Systems befinden. Und diese Darstellung ist auch gar nicht ganz falsch, wenn man sich dabei vorstellt, daß mentale Repräsentationen, die sich in der Wunschbox befinden, eine andere funktionale Rolle spielen als mentale Repräsentationen, die sich in der Überzeugungsbox befinden. (Man kann z. B. annehmen, daß das Verhalten des Systems u. a. durch folgendes Gesetz bestimmt wird: „Wenn sich in der Wunschbox des Systems eine mentale Repräsentation mit der Bedeutung, daß *p*, befindet und in der Überzeugungsbox eine mentale Repräsentation mit der Bedeutung, daß die Handlung *X* ein geeignetes Mittel zur Erreichung von *p* ist, dann führt das System die Handlung *X* aus.“) Aber die Annahme wirklicher Boxen ist natürlich nicht zwingend mit Fodors Theorie verbunden. Wichtig ist nur, daß die mentalen Repräsentationen in einem System verschiedene Rollen spielen können und daß diese Rollen den verschiedenen Typen von intentionalen Zuständen entsprechen.

b. Die Teile dieser Strukturen sind „transportierbar“; dieselben Teile können in verschiedenen Repräsentationen auftreten.

c. Mentale Repräsentationen haben als strukturierte Gebilde eine *kombinatorische Semantik*: die Art von Semantik, bei der die Bedeutung (relativ) komplexer Ausdrücke in regelhafter Weise durch die Bedeutung ihrer (relativ) einfachen Teile bestimmt wird (Fodor 1987, 135–137).

Wie hängen nun diese beiden „Implementierungsthesen“ (RTG) und (LT) mit den Thesen (I1)–(I5) zusammen? Welche Gründe sprechen Fodor zufolge für diese Thesen?

Ich hatte schon erwähnt, daß für Fodor aufgrund von (I1) die Annahme plausibel ist, daß auch die Realisierungen intentionaler Zustände relationalen Charakter haben bzw. daß den beiden Aspekten (Typ und Inhalt) intentionaler Zustände auch zwei Aspekte ihrer physischen Realisierung entsprechen. Diese Annahme wird für ihn sogar noch plausibler, wenn sie mit der weiteren Annahme verbunden ist, daß dem Inhaltsaspekt mentale Repräsentationen entsprechen, weil diese Annahme zusätzlich eine einfache Erklärung für (I2) liefert, d. h. für die semantische Bewertbarkeit intentionaler Zustände. Denn mentale Repräsentationen sind selbst Bedeutungsträger und vererben daher ihre Wahrheits- bzw. Erfüllungsbedingungen den intentionalen Zuständen, die mit ihrer Hilfe realisiert sind. Allerdings ist das Problem damit natürlich nur verschoben. Jetzt stellt sich die Frage, woher mentale Repräsentationen ihre Bedeutung bekommen. Aber immerhin führt die These (RTG) doch zu einer Vereinfachung in diesem Bereich. Fodor betont an vielen Stellen, daß zunächst einmal zwei Arten von „Dingen“ Bedeutung, d. h. einen repräsentationalen Inhalt zu haben scheinen: Symbole und intentionale Zustände (z. B. 1987, xi). Wenn jedoch intentionale Zustände im Sinne von (RTG) analysiert werden können, dann bleiben als eigentliche Bedeutungsträger nur Symbole und mentale Repräsentationen übrig, d. h. Entitäten von ungefähr derselben Art. Wenn (RTG) zutrifft, ergibt sich also eine Vereinheitlichung im Hinblick auf das Problem möglicher Bedeutungsträger.

Die Annahme, daß dem Inhaltsaspekt intentionaler Zustände mentale Repräsentationen entsprechen, liefert möglicherweise auch eine Erklärung für (I3), also für die Opakheit dieser Zustände. Fodor ist in diesem Punkt zwar nicht sehr explizit, aber folgende Argumentation scheint doch möglich: Wenn wir *O* die Überzeugung (Ü1) zuschreiben, indem wir sagen, *O* glaubt, daß Scott Scott ist, bzw. die Überzeugung (Ü2), indem wir sagen, *O* glaubt, daß Scott der Autor des „Waverly“ ist, dann sind diese Überzeugungen dadurch realisiert, daß *O* in der entsprechenden Relation *R* zu den mentalen Repräsentationen Σ bzw. Σ' steht. Weiter können wir – besonders auch aufgrund von (LT) – annehmen, daß Σ und Σ' aus Teilausdrücken „a“, „b“ und „ \equiv “ aufgebaut sind, die den Ausdrücken „Scott“, „Autor des ‚Waverly‘“ und „ist (identisch mit)“ entsprechen und für die insbesondere gilt, daß „a“ und „b“ zwar denselben Bezug, aber verschiedene Bedeutung haben. Weiter gilt, daß zwei Ausdrücke der lingua mentis genau dann bedeutungsgleich

sind, wenn sie dieselbe funktionale/computationale Rolle spielen.⁸ Also haben „a“ und „b“ und daher auch Σ und Σ' verschiedene funktionale/computationale Rollen. Die Zustände, die darin bestehen, daß O in der Relation R zu Σ bzw. Σ' steht, müssen also verschieden sein, da sie aufgrund der verschiedenen funktionalen/computationalen Rollen von Σ und Σ' verschiedene kausale Kräfte, d. h. verschiedene Ursachen und Wirkungen haben.⁹ Wären „a“ und „b“ dagegen nicht nur bezugs-, sondern auch bedeutungsgleich, dann wären dementsprechend (Ü1) und (Ü2) Überzeugungen derselben Art.

Der Zusammenhang zwischen (LT) und (I4) schließlich ist ziemlich offenkundig. Dennoch denke ich, daß die bisherigen Überlegungen nicht den Kern der Fodorschen Argumentation ausmachen. Denn entscheidend für diese Überlegungen scheint mir letzten Endes die Beziehung zwischen (RTG) und (LT) und den Thesen (I5) bzw. (I5') zu sein.

Für Fodor sind nämlich die Kausalbeziehungen, die zwischen intentionalen Zuständen bestehen, Kausalbeziehungen zwischen höherstufigen Eigenschaften, die im Rahmen der Spezialwissenschaft Psychologie untersucht werden, und solche Kausalbeziehungen beruhen seiner Meinung nach immer auf physischen Mechanismen, die die physischen Realisierungen der entsprechenden Zustände kausal miteinander verbinden (1989, 155). Wenn den Kausalbeziehungen zwischen intentionalen Zuständen in vielen Fällen Rationalitätsprinzipien entsprechen, dann stellt sich für Fodor daher unabweisbar die Frage, wie physische *Mechanismen* denn überhaupt aussehen können, die solchen Kausalbeziehungen zugrunde liegen, d. h. wie physische Mechanismen aussehen können, die ihrerseits Rationalitätsprinzipien parallelisieren (1987, 14). Soweit ich sehen kann, ist dies genau die Frage, die Fodor dazu führt, auf die Idee der Symbolverarbeitung zurückzugreifen. Und das ist alles andere als überraschend. Denn auf der einen Seite hat die Beweistheorie, auf die sich Fodor immer wieder bezieht, gezeigt, daß der Begriff der logischen Folgerung formalisiert werden kann, d. h. daß man diesen Begriff durch Bezugnahme auf geeignete Kalküle auch rein syntaktisch charakterisieren kann. Und auf der anderen Seite haben die Computerwissenschaften gezeigt, daß sich solche syntaktische Umformungsprozesse mit Hilfe von Symbolverarbeitungsprozessen physisch realisieren lassen. Mit anderen Worten: Physische Mechanismen, die Rationalitätsanforderungen respektieren, lassen sich mit Symbolverbreitungsprozessen realisieren. Und was für Fodor mindestens ebenso wichtig ist: Es ist nicht zu sehen, welche anderen Mechanismen diese Aufgabe übernehmen könnten.¹⁰

Damit wird aber auch ein weiterer Grund dafür klar, warum für Fodor (RTG) nicht genügt, warum zu dieser

These die These (LT) hinzukommen muß. Symbolverarbeitungsprozesse setzen strukturierte Repräsentationen voraus: sie bestehen geradezu darin, daß sie Repräsentationen nach Regeln verändern, die auf die Struktur dieser Repräsentationen Bezug nehmen. Wenn kausale Mechanismen, die Rationalitätsanforderungen respektieren, nur auf Symbolverarbeitungsprozessen beruhen können und wenn Symbolverarbeitungsprozesse strukturierte Repräsentationen voraussetzen, ist der Schluß auf (LT) also offenbar unvermeidlich.

Allerdings, und damit kommen wir zu einem letzten entscheidenden Punkt der Fodorschen Überlegungen, gilt dies nur unter bestimmten Voraussetzungen. Nehmen wir einmal an, daß ein bestimmtes System S nur endlich viele Überzeugungen haben kann und daß es daher – (RTG) zufolge – auch nur endlich viele mentale Repräsentationen gibt, zu denen S in der entsprechenden Relation R stehen kann. Dann läßt sich das Problem der „rationalen“ Verursachung, wie ich es einmal nennen möchte, sehr viel einfacher lösen. Denn in diesem Fall reicht es, wenn das System für jeden der nur endlich vielen Fälle, in denen es rational ist, p_{n+1} zu glauben, wenn man p_1, \dots, p_n glaubt, über einen eigenen Mechanismus verfügt, der dafür sorgt, daß die p_{n+1} entsprechende mentale Repräsentation $mr(p_{n+1})$ in der Überzeugungs-Box von S abgespeichert wird, wenn sich $mr(p_1), \dots, mr(p_n)$ schon in dieser Box befinden. Für Fodor ist deshalb ein weiteres Merkmal intentionaler Zustände von entscheidender Bedeutung, das so formuliert werden kann:

(I6) Intentionale Zustände sind produktiv¹¹ und systematisch.

Daß intentionale Zustände produktiv sind, soll in diesem Zusammenhang heißen, daß jedes System potentiell z. B. unendlich viele Überzeugungen haben kann. Wenn Sp glaubt, kann es auch $\neg p$ glauben und $p \vee q$ und $p \vee r$ usw. Daß intentionale Zustände systematisch sind, heißt dagegen, daß wenn Sp glaubt, S auch in der Lage ist, Dinge zu glauben, die mit p systematisch zusammenhängen. D. h. wenn S glaubt, daß Hans Gerda liebt, muß S auch glauben können, daß Gerda Hans liebt oder daß Hans Helga liebt usw.

Für Fodor ist nun zweierlei wichtig. Erstens, daß die Produktivität und Systematizität intentionaler Zustände selbst einer Erklärung bedarf und daß seiner Meinung nach (LT) die beste Erklärung dieser Art liefert. Und zweitens: Wenn intentionale Zustände tatsächlich produktiv und systematisch sind, dann scheint es keine Alternative zu der Annahme zu geben, daß die kausalen Relationen zwischen intentionalen Zuständen, denen Rationalitätsprinzipien entsprechen, auf beweistheoretisch fundierten, struktursensitiven Symbolverarbeitungsprozessen beruhen.

Ich denke, daß damit zumindest in groben Zügen klar geworden ist, wie für Fodor die zentralen Thesen (RTG) und (LT) seiner repräsentationalen Theorie mit seiner Analyse der charakteristischen Merkmale intentionaler Zustände zusammenhängen.

⁸ Dies ist in diesem Zusammenhang natürlich die entscheidende Prämisse.

⁹ Zur Rolle von kausalen Kräften bei der Typ-Identifikation von Zuständen vgl. auch Abschnitt 2.2.

¹⁰ Genau an dieser Stelle setzt die Kritik der Konnektionisten am Symbolverarbeitungsansatz ein. Denn Netzwerke scheinen, was von Fodor allerdings aus Gründen, die mit dem Prinzip (I6) (s. u.) zu tun haben, bestritten wird, eine alternative Realisierungsmöglichkeit darzustellen.

¹¹ In (1987, 148) rückt Fodor von der Überzeugung, daß intentionale Zustände produktiv sind, allerdings zum Teil wieder ab, so daß als entscheidendes Merkmal die Systematizität bleibt.

2. Ich komme zum zweiten Teil der Kemmerlingschen Überlegungen, seiner grundsätzlichen Kritik der Theorie Fodors. Dabei will ich nicht auf die Kritikpunkte eingehen, die Kemmerling nur kurz erwähnt, die aber nicht den Kern seiner Überlegungen ausmachen, da sie im wesentlichen schon von anderen Autoren vorgebracht wurden: die Möglichkeit nicht-expliziter Repräsentationen, das Frame-Problem, den Konnektionismus als Alternative.

Kemmerlings eigene Kritik bezieht sich im wesentlichen auf Fodors Antwort auf die Frage, wie die Bedeutung mentaler Repräsentationen in naturalistischer Weise spezifiziert werden kann. Dabei kommt er zu dem Ergebnis: 1. Fodors Theorie zum weiten Inhalt intentionaler Zustände ist falsch; 2. Fodors Theorie zum engen Inhalt intentionaler Zustände ist „dubios bis hin zu der Frage, ob es dabei wirklich um Semantisches geht“ (57 a). Beide Punkte zusammen legen seiner Meinung nach den Schluß nahe, daß der ganze Repräsentationalismus verfehlt ist.

Ich möchte im folgenden sowohl zu den Punkten 1. und 2. als auch zu der Frage Stellung nehmen, ob 1. und 2. wirklich die von Kemmerling gezogene Schlußfolgerung plausibel machen.

2.1. *Weiter Inhalt.* Soweit ich sehen kann, stellt Kemmerling im Abschnitt 3 seines Artikels Fodors Überlegungen zur Frage nach der Bedeutung mentaler Repräsentationen völlig korrekt dar. Seine Kritik an diesen Überlegungen erscheint mir jedoch nicht so überzeugend. Fodors Theorie läßt sich kurz so zusammenfassen:

- (BED) Ein Ausdruck μ der *lingua mentis* hat genau dann die Bedeutung x , wenn die folgenden beiden Bedingungen erfüllt sind:
- unter psychophysisch optimalen Bedingungen verursachen alle x μ ;
 - falls (in einigen Fällen) μ durch y verursacht wird, hängt die Kausalbeziehung zwischen y und μ asymmetrisch von der Kausalbeziehung zwischen x und μ ab.

Kemmerling zufolge hat diese Theorie zwei Schwachpunkte. Zum einen läßt sich seiner Meinung nach der Ausdruck „psychophysisch optimale Bedingungen“ nicht naturalistisch explizieren, und zum anderen leistet der Begriff asymmetrische Abhängigkeit nicht das, was er leisten soll.

Zum ersten Punkt möchte ich nur kurz anführen, daß Kemmerling die entscheidenden Argumente schuldig bleibt. Soweit ich ihn verstanden habe, laufen seine Überlegungen zu diesem Punkt darauf hinaus, daß der Begriff „psychophysisch optimale Bedingungen“ nur durch Bezugnahme auf den Begriff „intakter Beobachter“ expliziert werden kann und daß dieser Begriff eine teleologische Komponente enthält. Doch dies macht den Schluß, daß der Ausdruck „psychophysisch optimale Bedingungen“ nicht naturalistisch expliziert werden kann, nur plausibel, wenn gezeigt wird, daß sich diese Komponente ihrerseits einer naturalistischen Analyse entzieht. Und dies zeigt Kemmerling nicht, er behauptet es nur.¹²

¹² Allerdings könnte Kemmerling mit Hinweis auf die Literatur zu dieser Frage wohl zu Recht argumentieren, daß die Beweislast in Bezug auf diesen Punkt offensichtlich auf der Seite Fodors liegt.

Aber zum zweiten Punkt, der wichtiger ist: Es ist sicher so, daß Fodor nur sehr ungenügend klar macht, was er in diesem Zusammenhang unter asymmetrischer Abhängigkeit verstanden wissen will. Aber daraus allein folgt noch nicht, daß sich dies auch nicht klar machen läßt. Kemmerling argumentiert, daß diese Idee keinen Sinn macht, da sich unter den verschiedenen Ursachen, die μ hervorrufen, keine entsprechende Unterscheidung treffen läßt. Wenn y 'e unter bestimmten Umständen μ hervorrufen, dann werden andere y 'e unter naturalistisch gesehen gleichen Umständen wieder μ hervorrufen, ganz egal, welche andere Ursachen μ unter anderen Umständen haben kann. Insofern, so Kemmerling, ist die Kausalextension von μ homogen (S.55 a). Das ist plausibel. Aber mit dieser Argumentation läßt Kemmerling andere Wege völlig außer acht, die man bei dem Versuch, den Begriff asymmetrische Abhängigkeit von Kausalbeziehungen zu explizieren, vielleicht auch gehen kann.

Wahrscheinlich ist doch, daß sich der Mechanismus M , der in O dafür sorgt, daß x 'e bzw. y 'e zu μ führen, selbst ontogenetisch oder phylogenetisch entwickelt hat. Wenn das so ist, ist es aber durchaus möglich, daß M in O entstanden ist, weil x 'e für O interessant sind und weil es deshalb für O wichtig ist, über einen Mechanismus zu verfügen, der dafür sorgt, daß x 'e bestimmte interne Repräsentationen verursachen. Und dann ist es auch möglich (oder sogar wahrscheinlich), daß sich in O zu diesem Zweck ein nicht vollständig verlässlicher Mechanismus entwickelt, d. h. ein Mechanismus, der es auch erlaubt, daß manchmal y 'e μ hervorrufen, obwohl y 'e für O ganz ohne Interesse sind. In einem solchen Fall scheint es mir plausibel anzunehmen, daß sich M in O auch dann entwickelt hätte, wenn es in der Welt zwar x 'e, aber keine y 'e gegeben hätte, daß das Umgekehrte aber nicht der Fall ist. Und dies würde, wie mir scheint, bedeuten, daß wir es hier tatsächlich mit einem Fall zu tun haben, in dem man zumindest mit einer gewissen Plausibilität sagen kann, daß die Kausalbeziehung zwischen y und μ asymmetrisch von der zwischen x und μ abhängt.¹³

¹³ Gegen diese Argumentation könnte man einwenden, daß sie nur deshalb zum Erfolg führe, weil sie sich auf teleologische Überlegungen (z. B. im Hinblick darauf, was für ein System „wichtig“ oder „interessant“ ist) stütze und damit den von Fodor vorgegebenen Rahmen eindeutig überschreite. Denn Fodor schreibe ausdrücklich, daß sein Programm nur dann erfolgreich sei, wenn es gelinge, Bedingungen für die Bedeutung mentaler Repräsentationen in nicht-intentionalem, nicht-semantischem und nicht-teleologischem Vokabular zu formulieren (1987, 126).

Dieser Einwand wäre jedoch nicht berechtigt. Denn die teleologischen Aspekte der vorgetragenen Argumentation dienen nur der Veranschaulichung. In der Definition der asymmetrischen Abhängigkeit von Kausalbeziehungen, auf die diese Argumentation hinausläuft, spielen sie überhaupt keine Rolle, denn diese Definition könnte man so formulieren:

(AAK) Die Kausalbeziehung zwischen y und μ hängt asymmetrisch von der Kausalbeziehung zwischen x und μ ab, wenn der Mechanismus M , auf dem diese Kausalbeziehungen beruhen, in O auch dann entstanden wäre, wenn es in O 's Umgebung keine y 'e, aber x 'e gegeben hätte, wenn dieser Mechanismus in O aber nicht entstanden wäre, wenn es in O 's Umgebung keine x 'e, aber y 'e gegeben hätte.

Es geht mir hier jedoch gar nicht darum, Fodors Versuch, die Bedeutung mentaler Repräsentationen naturalistisch zu explizieren, zu verteidigen. Was ich zeigen will, ist nur, daß Fodors Ansatz keineswegs so offensichtlich verfehlt ist, wie Kemmerling zu glauben scheint.

2.2. *Enger Inhalt.* Neben Fodors Überlegungen zum weiten Inhalt mentaler Repräsentationen sind es besonders dessen entsprechende Überlegungen zum engen Inhalt, die Kemmerling für völlig „dubios“ hält. Dabei ist sein Argument nicht, daß sich dieser Begriff nicht vernünftig explizieren ließe. Dies geht auch Kemmerling zufolge durchaus: Der enge Inhalt einer mentalen Repräsentation bzw. eines Ausdrucks der *lingua mentis* ist eine Funktion, die dieser Repräsentation bzw. diesem Ausdruck für jeden möglichen Kontext eine Bedeutung (einen weiten Inhalt) zuordnet. Kemmerling zufolge sind Fodors Überlegungen zu diesem Punkt aber mit zwei anderen Problemen belastet. Zum einen finde sich bei Fodor überhaupt keine Antwort auf die Frage, welche der vielen möglichen Funktionen denn nun tatsächlich den engen Inhalt eines bestimmten Ausdrucks ausmache. Und zum anderen sei nicht einmal sicher, daß bei diesem ganzen Ansatz etwas herauskomme, was überhaupt noch den Namen „semantisch“ verdiene. Dies alles will ich gerne zugestehen, weil ich ohnehin glaube, daß Fodors Suche nach dem engen Inhalt ganz und gar fehlgeleitet ist. Doch was folgt daraus für Fodors repräsentationalistischen Ansatz im allgemeinen? Mir scheint, nicht viel. Diese Antwort wird jedoch vielleicht nur dann plausibel, wenn man sich den Kontext der Fodorschen Überlegungen zum engen Inhalt noch einmal vergegenwärtigt.

Soweit ich sehen kann, beruhen diese Überlegungen im wesentlichen auf drei Thesen, einer allgemein wissenschaftstheoretischen, einer, die sich speziell auf die Wissenschaftstheorie der intentionalen Psychologie bezieht, und schließlich einer eher ontologischen These. Die allgemeine wissenschaftstheoretische These lautet:¹⁴

(T1) In respektablen Wissenschaften dürfen bei der Formulierung von Gesetzen nur solche Begriffe verwendet werden, für die gilt, daß zwei Gegenstände *x* und *y* genau dann unter sie fallen, wenn sie über dieselben kausalen Kräfte verfügen.

Die zweite These ist eine Variante der These (I1):

(T2) Die intentionale Psychologie unterscheidet intentionale Zustände nur nach zwei Aspekten: ihrem Typ und ihrem Inhalt nach.

Die ontologische These schließlich lautet:

(T3) Die kausalen Kräfte eines Gegenstands supervenieren auf seinen physischen Eigenschaften (Fodor 1978, 30).

Es ist leicht zu sehen, wie sich aus diesen drei Thesen das Motiv für Fodors Suche nach dem engen Inhalt ergibt. Offensichtlich verwendet die intentionale Psychologie bei der Formulierung ihrer Gesetze Begriffe, die sich auf intentionale Zustände beziehen. Nach (T2) werden solche Zustände aber nur im Hinblick auf zwei Aspekte unter-

schieden: ihrem Typ und ihrem Inhalt nach, d. h. zwei intentionale Zustände sind genau dann Zustände derselben Art, wenn sie sich weder dem Typ noch dem Inhalt nach unterscheiden. Wenn die Psychologie eine respektable Wissenschaft sein will, gehören – nach (T1) – intentionale Zustände aber genau dann zur selben Art, wenn sie dieselben kausalen Kräfte haben. Also gilt: Wenn die Psychologie eine respektable Wissenschaft sein will, müssen intentionale Zustände, die sich weder dem Typ noch dem Inhalt nach unterscheiden, dieselben kausalen Kräfte haben. Aus (T3) ergibt sich aber weiter, daß alle Zustände zweier Gegenstände, die sich physisch nicht unterscheiden, dieselben Kräfte haben. Also folgt, daß sich diese beiden Gegenstände auch in denselben intentionalen Zuständen befinden.

Nun zeigen die Beispiele von Putnam und Burge aber genau, daß Situationen denkbar sind, in denen wir geneigt sind, zwei Personen, die sich physisch nicht unterscheiden, dennoch Überzeugungen mit verschiedenem Inhalt zuzuschreiben. *O* glaubt, daß Aluminium ein Leichtmetall ist, während *O*'s Doppelgänger auf der Zwillingserde glaubt, daß Zwalum ein Leichtmetall ist.

Fodor zufolge zeigt dies jedoch nur eins: daß wir bei der normalen Praxis der Bestimmung des Inhalts von intentionalen Zuständen Kriterien verwenden, die mit den in der These (T1) formulierten Anforderungen an die Begriffsbildung in respektablen Wissenschaften nicht vereinbar sind (Kriterien des weiten Inhalts). Da intentionale Zustände jedoch auch in der wissenschaftlichen Psychologie nur nach Typ und Inhalt unterschieden werden, muß die wissenschaftliche Psychologie andere Kriterien zur Bestimmung des Inhalts intentionaler Zustände verwenden (Kriterien des engen Inhalts).

Soweit der Versuch, Fodors Motive zu rekonstruieren. Was haben nun aber diese Überlegungen mit den Thesen (RTG) und (LT) zu tun, die den Kern seiner Repräsentationalen Theorie des Geistes ausmachen? Meines Erachtens ist die Antwort ganz klar: Nichts. Fodors Überlegungen zum engen Inhalt stehen und fallen mit der allgemeinen These (T1), die von seinen sonstigen Annahmen völlig unabhängig ist. Wenn sich diese These als falsch erwiese,¹⁵ hätte dies daher keinerlei Folgen, weder für Fodors in den Thesen (RT) und (I1)–(I6) zusammengefaßte Analyse der intentionalen Psychologie noch für die Überlegungen, die Fodor von dieser Analyse zu den Thesen (RTG) und (LT) führen. Mit anderen Worten: Fodors Thesen zum engen Inhalt stehen orthogonal zu seinen Thesen zur Repräsentationalen Theorie des Geistes. Auch wenn Kemmerling mit seiner Kritik an Fodors Theorie des engen Inhalts recht hat, folgt daraus also nichts für die Repräsentationale Theorie des Geistes im allgemeinen.

2.3. *Schlußfolgerungen.* Damit ist auch schon die Frage zum Teil beantwortet, ob Kemmerlings Kritik an der Bedeutungskomponente der Fodorschen Theorie tatsächlich den weitreichenden Schluß zuläßt, daß die ganze Repräsentationale Theorie des Geistes auf schwachen

¹⁴ Vgl. Kemmerling, 50b.

¹⁵ Und sie ist wahrscheinlich falsch. Vgl. z. B. Egan (1990).

Füßen steht. Für Kemmerlings Kritik an Fodors Thesen zum engen Inhalt gilt das sicher nicht. Aber wie sieht die Sache für seine Kritik an Fodors Thesen zum weiten Inhalt aus?

Ich hatte im Abschnitt 2.1 schon betont, daß mir diese Kritik etwas zu kurz zu greifen scheint. Aber lassen wir das einmal beiseite. Was würde folgen, wenn Kemmerling mit seiner Kritik in diesem Punkt recht hätte?

Nun, die These (RTG) wäre sicher in Schwierigkeiten, wenn sich herausstellen sollte, daß es prinzipiell unmöglich ist, plausibel zu erläutern, unter welchen Bedingungen eine mentale Repräsentation *mr* welche Bedeutung hat.¹⁶ Aber dies ist natürlich nicht schon dann gezeigt, wenn man gezeigt hat, daß Fodors eigene Antwort auf diese Frage nicht plausibel ist. Kemmerling behauptet zwar (S. 55 b), Fodors Theorie sei „das beste Stück in diesem Regal“. An dieser Behauptung sind jedoch einige Zweifel erlaubt, zumal Kemmerling als mögliche Alternativen nur die Kausaltheorien von Dretske und Stampe explizit erwähnt. Zum einen nämlich ist bis jetzt noch keineswegs ausgemacht, daß sich diese Theorien nicht doch noch so verbessern lassen, daß sie eine plausible Antwort auf das Disjunktionsproblem (vgl. Kemmerling 51 b–52 a) ermöglichen, und zum anderen gibt es noch weitere Alternativen, z. B. die funktional-teleologische Theorie von Millikan und Papineau und den interpretationstheoretischen Ansatz von Haugeland und Cummins. Ich persönlich denke, daß der letztgenannte Ansatz die besten Chancen hat, zu einer befriedigenden Antwort zu führen. Und dies insbesondere deshalb, weil er – anders als alle anderen mir bekannten Theorien – nicht davon ausgeht, daß es eine Frage der Tatsachen ist, welche Bedeutung mentale Repräsentationen haben.¹⁷ Nach Haugeland und Cummins ist es eher eine Frage der Interpretation (und vorab der Interpretierbarkeit), welche Bedeutung man einer mentalen Repräsentation zuschreibt. Und da sind durchaus verschiedene Interpretationen möglich.

Aber wie dem auch sei. Mir ging es nur darum zu zeigen, daß eine Kritik der Fodorschen Bedeutungstheorie allein sicher nicht ausreicht, um die These zu begründen, daß es überhaupt keine plausible Bedeutungstheorie für mentale Repräsentationen gibt. Es gibt Alternativen zu dieser Theorie, und es gibt sogar Alternativen, die viel plausibler sind als Fodors Ansatz. Insofern greift Kemmerlings Kritik viel zu kurz. Seine Argumente reichen sicher nicht aus, um den Repräsentationalismus zu Fall zu bringen. Und dies gilt um so mehr, als er auf die entscheidenden Argumente, die Fodor zufolge *für* diese Theorie sprechen (die Produktivität und Systematizität intentionaler Zustände und die Tatsache, daß zwischen intentionalen Zuständen Kausalbeziehungen bestehen, die semantische bzw. Rationalitätsbeziehungen respektieren), überhaupt nicht eingeht.

Kemmerling schließt seinen Artikel mit dem Fazit: „Das Gesagte beweist gewiß nicht, aber spricht doch ent-

schieden für das Ergebnis: Mentale Repräsentation findet ohne mentale Repräsentationen statt“ (57 a). Dies ist, wie mir scheint, ein schlichtes *non sequitur*, zu dem Kemmerling nur deshalb kommt, weil er davon ausgeht, daß Fodors gesamte Theorie unterminiert ist, wenn es gelingt, auch nur einen Stein aus ihr herauszubringen. Diese Einschätzung ist meiner Meinung nach jedoch durch nichts begründet. Den Kern der Repräsentationalen Theorie des Geistes bilden die Thesen (RTG) und (LT). Und diese Thesen sind unabhängig von den übrigen Thesen Fodors, d. h. jemand, der die Thesen (RTG) und (LT) vertritt, ist damit *nicht* darauf festgelegt, zugleich die These (BED) oder Fodors Thesen zum engen Inhalt zu vertreten. Wie schon gesagt: Natürlich wäre der Repräsentationalismus in Schwierigkeiten, wenn *jeder* Versuch scheitern sollte, eine Bedeutungstheorie für mentale Repräsentationen zu entwickeln. Aber er ist nicht auf Fodors Bedeutungstheorie und schon gar nicht auf Fodors Thesen zum engen Inhalt festgelegt. Selbst wenn Kemmerling gezeigt hätte, daß diese beiden Theoriestücke unhaltbar sind, hätte das daher für die Plausibilität der Kernthesen (RTG) und (LT) noch keinerlei Folgen.

3. Zum Schluß nur noch eine kurze Bemerkung. Es ging mir in dieser Replik auf Kemmerlings Artikel „Mentale Repräsentationen“ nur darum zu zeigen, daß die Kernthesen der Repräsentationalen Theorie des Geistes durch eine Kritik an Fodors Thesen zur Bedeutung mentaler Repräsentationen nicht wirklich berührt sind. Es ging mir nicht darum, diese Thesen bzw. Fodors Version dieser Thesen zu verteidigen. Ich denke, daß diese Version eine ganze Reihe von schwachen Punkten hat, z. B. die Grundthese (PRT) oder auch die in (RTG) enthaltene völlig überzogene Forderung, daß alle Inhalte *explizit* repräsentiert sein müssen. Ich habe jedoch den Eindruck gewonnen, daß es sich manche Kritiker der Fodorschen Theorie zu leicht machen. So einfach, wie es häufig dargestellt wird, ist diese Bastion der Kognitionswissenschaft nicht einzunehmen. Und bei aller Kritik an Fodor: Man sollte nicht übersehen, daß Fodor immerhin versucht, eine theoretisch äußerst interessante Frage zu beantworten, die Frage nämlich, ob sich aus der Tatsache, daß man das Verhalten eines Systems auf intentionale Weise verlässlich erklären und vorhersagen kann, allgemeine Aussagen über die interne Struktur dieses Systems ableiten lassen.¹⁸

Literatur

- Burge, T. (1979). Individualism and the Mental. *Midwest Studies in Philosophy*, 4, 73–121
 Cummins, R. (1989). *Meaning and Mental Representation*. Cambridge, Mass.: MIT Press
 Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge, Mass.: MIT Press

¹⁶ Es scheint mir allerdings nicht ganz klar zu sein, ob es für (RTG) auch schon katastrophal wäre, wenn sich nur herausstellen sollte, daß dies nicht in naturalistischen Termini möglich ist.

¹⁷ Insofern beruht dieser Ansatz offenkundig auf entsprechenden Überlegungen von Quine und Davidson.

¹⁸ Noch eine kurze Bemerkung: Ich bin bei meiner Darstellung der Fodorschen Theorie an einigen Stellen von Fodors eigenen Formulierungen abgewichen. Diese Abweichungen scheinen mir aber durch das allgemeine Prinzip der philosophischen Interpretation von Texten legitimiert, daß man dem interpretierten Autor – manchmal auch gegen den Wortlauf seiner Texte – eine möglichst kohärente und plausible Theorie zuschreiben sollte.

A. Beckermann: Endgültiger Todesstoß für den Repräsentationalismus?

- Dretske, F.** (1986). Misrepresentation. In: R. J. Bogdan (ed.), *Belief*, pp. 17–36. Oxford: Oxford University Press
- Egon, M.F. (1990). Must Psychology be Individualistic? Report Nr. 25/1990 der Forschungsgruppe „Mind and Brain“ des ZiF, Universität Bielefeld
- Fodor, J. A. (1975). *Language of Thought*. New York: Crowell
- Fodor, J. A. (1978). Propositional Attitudes. *The Monist*, 64, 501–523. Wiederabgedr. in Fodor (1981), 177–203
- Fodor, J. A. (1981). *Representations*. Cambridge, Mass.: MIT Press
- Fodor, J. A. (1985). Fodor's Guide to Mental Representation. *Mind*, 94, 76–100. Wiederabgedr. in Fodor (1990), 3–29
- Fodor, J. A. (1987). *Psychosemantics*. Cambridge, Mass.: MIT Press
- Fodor, J. A. (1989). Making Mind Matter More. *Philosophical Topics*, 17, 59–79. Wiederabgedr. in Fodor (1990), 137–159
- Fodor, J. A. (1990). *A Theory of Content and Other Essays*. Cambridge, Mass.: MIT Press
- Haugeland, J. (ed.) (1981 a). *Mind Design*. Cambridge, Mass.: MIT Press
- Haugeland, J. (1981 b). Semantic Engines. In: Haugeland (1981 a), 1–34
- Haugeland, J. (1985). *Artificial Intelligence. The Very Idea*. Cambridge, Mass.: MIT Press
- Kemmerling, A. (1990). Mentale Repräsentationen. *Kognitionswissenschaft*, 1, 47–57
- Millikan, R. (1984). *Language, Thought, and Other Biological Categories*. Cambridge, Mass.: MIT Press
- Millikan, R. (1989). Biosemantics. *Journal of Philosophy*, 86, 281–297
- Papineau, D. (1988). *Reality and Representation*. Oxford: Basil Blackwell
- Putnam, H. (1975). The Meaning of 'Meaning'. In: H. Putnam. *Mind, Language and Reality. Philosophical Papers. Vol. 2*, pp. 215–271. Cambridge: Cambridge University Press
- Putnam, H. (1988). *Representation and Reality*. Cambridge, Mass.: MIT Press
- Stampe, D. (1977). Towards a Causal Theory of Linguistic Representation. *Midwest Studies in Philosophy*, 2, 42–63