

Mentale Verursachung und metaphysischer Realismus

VON GODEHARD BRÜNTRUP S.J.¹

Eine „Wende“ in der analytischen Philosophie

H. Putnam hat die These aufgestellt, daß die analytische Philosophie sich von einer in ihren Anfängen anti-metaphysischen Bewegung zur am meisten metaphysisch orientierten philosophischen Strömung der Gegenwart entwickelt habe². Nirgendwo läßt sich diese Entwicklung besser ablesen als in der Philosophie des Geistes, vor allem in der Kontroverse um das Problem der Möglichkeit mentaler Verursachung. Die zentrale Frage lautet hier: Wie ist es auf dem Hintergrund einer physikalistischen Ontologie möglich, mentalen Ereignissen oder Eigenschaften kausale Wirksamkeit zuzusprechen? Nicht wenige kontinentaleuropäische Autoren sahen sich gerade durch diese Wiederbelebung klassischer metaphysischer Problemstellungen gerechtfertigt, die entsprechenden Debatten in der analytischen Philosophie für schulinterne Probleme in einer historisch unkritischen Denkweise zu halten³. Auch innerhalb der analytischen Tradition waren Stimmen zu hören, die forderten, die metaphysische Denkweise fallenzulassen, um so das unlösbare Problem der mentalen Verursachung zu entschärfen⁴. Im folgenden geht es mir darum zu zeigen, an genau welchen Knotenpunkten der aktuellen Debatte der Eindruck der „Unlösbarkeit“ des Problems mentaler Verursachung von ganz bestimmten (wenn auch oft impliziten) metaphysischen Hintergrundannahmen provoziert wird. Weiter will ich zeigen, in welchem Sinne die Ablehnung des metaphysischen Realismus sich gerade an dieser Stelle als Ausweg aufdrängt⁵.

Das Grundproblem in einer Standardformulierung

Das Problem der mentalen Verursachung läßt sich am besten auf dem Hintergrund einer allgemeinen Formulierung des Leib-Seele-Problems dar-

¹ Für kritische Kommentare zu einer früheren Fassung dieses Textes danke ich Peter Bieri, Andreas Kemmerling, Brian McLaughlin, Piotr Lenartowicz, Lorenz Bruno Puntel, Edmund Runggaldier und Raimo Tuomela. Alle verbliebenen Unklarheiten und Fehler gehen selbstverständlich allein zu meinen Lasten.

² H. Putnam, *Renewing Philosophy*, Cambridge, MA 1993, 187.

³ Vgl. z. B. *Habermas*, „Der analytische Materialismus hat mich tatsächlich nie sonderlich beeindruckt, eben weil er eine metaphysische Position ist...“ *J. Habermas*, *Nachmetaphysisches Denken*, Frankfurt am Main 1988, 29.

⁴ Vgl. z. B. B. Baker: „Start with successful explanatory practices and let the metaphysics go. At least, we can avoid the insoluble problem of mental causation.“ *L. Baker*, *Metaphysics and Mental Content*, in: *J. Heil/A. Mele* (eds.), *Mental Causation*, Oxford 1993, 75–96, hier 95.

⁵ Ich kann den Argumentationsgang hier nur sehr prägnant darstellen. In meinem Buch *Mentale Verursachung – Eine Theorie aus der Perspektive des semantischen Anti-Realismus* (Stuttgart 1994) sind die kritische Auseinandersetzung und auch die Ausarbeitung einer Alternative in einer vollständigeren und technisch expliziteren Form entwickelt worden.

stellen. Folgende drei Prinzipien können nicht ohne Widerspruch gleichzeitig für wahr gehalten werden:

- [1] Die physische Welt ist kausal lückenlos geschlossen.
- [2] Aus der kausalen Geschlossenheit der physischen Welt folgt die kausale Wirkungslosigkeit mentaler Ereignisse oder Eigenschaften.
- [3] Die mentalen Ereignisse oder Eigenschaften sind kausal wirksam.

Die verschiedenen Lösungsversuche des Leib-Seele-Problems lassen sich als die Negation jeweils eines dieser drei Prinzipien klassifizieren. Die Negation von Prinzip [2] ist aus folgenden Gründen philosophisch besonders attraktiv:

- Man vermeidet die Schwierigkeiten des Dualismus (sich kreuzende Kausalketten)
- Man kann Physikalist bleiben, ohne die unerwünschten Konsequenzen des Epiphänomenalismus, Instrumentalismus oder Eliminativismus akzeptieren zu müssen (kausale Wirkungslosigkeit des Mentalen).
- Man kann die Alltagsüberzeugungen über die kausale Rolle unseres mentalen Lebens in eine wissenschaftlich orientierte Ontologie integrieren.

Die fundamentale metaphysische Intuition, die mit der Negation von Prinzip [2] verträglich zu sein scheint, ist die einer Schichtontologie. Im Gegensatz zum Dualismus mit seiner kaum intelligiblen Interaktion völlig verschiedener Substanzen versucht man das Geistige als eine höhere Organisationsstufe des Physischen zu begreifen. Die partikulären physischen Entitäten und ihre mereologischen Kompositionen sind alles, was es gibt. Die ungeheure Vielfalt der Ganzen, die aus diesen Teilen entstehen, läßt sich in verschiedene Stufen untergliedern. Jeder Stufe entspricht eine spezielle Wissenschaft (von Chemie über Psychologie zur Soziologie). Auf den ersten Blick scheint dieser ontologische Ansatz das Grundproblem des Dualismus, die Unerklärbarkeit mentaler Verursachung, vermeiden zu können. Auch ohne die Einführung nichtphysischer Substanzen scheint jeder Ebene eine gewisse Autonomie zuzukommen. Aber dieser Eindruck ist trügerisch. Ohne jeden Zweifel nimmt jede Ebene epistemisch-explanatorisch eine eigene Rolle ein, die sich sogar in der faktischen Irreduzibilität der Theorien der höherstufigen Wissenschaften auf die Theorien der basalen Wissenschaften ausdrücken könnte. Die Voraussetzung war aber, daß es sich um eine metaphysische Theorie handelt. Es geht um mehr als nur begriffliche Autonomie, es geht um eine ontologische Eigenständigkeit der von den speziellen Wissenschaften eingeführten Entitäten, so daß diese Entitäten in der Metaphysik der Kausalordnung der aktuellen Welt eine spezifische Rolle übernehmen.

Metaphysischer Realismus

Was genau heißt es, daß die Positionen, die Prinzip [2] negieren, in diesem Sinn metaphysische Theorien sind? Im vorliegenden Zusammenhang bedeutet das primär, daß die in ihnen behaupteten Aussagen im Rahmen einer

realistischen Semantik interpretiert werden. Präziser: Unter einer metaphysischen Position verstehe ich eine im Sinne Putnams metaphysisch-realistische Theorie mit folgenden (oder sehr ähnlichen) Grundannahmen:

– Die von uns unabhängige, objektive Realität macht jeden sinnvollen Behauptungssatz über diese Realität eindeutig entweder wahr oder falsch. Es gilt also das Prinzip der Bivalenz. Die Semantik ist zweiwertig.

– Die Welt besteht aus Entitäten, die unabhängig von unserem Wissen über sie existieren.

– Es gibt eine einzige wahre Theorie, welche die Welt so beschreibt, wie sie ist. Leider liegt diese Theorie außerhalb unserer epistemischen Reichweite (Gottesstandpunkt).

– Wahrheit kann nicht epistemisch verstanden werden. Selbst eine aus unserer Sicht ideale Theorie könnte immer noch falsch sein. Wahrheit muß metaphysisch (= non-epistemisch) verstanden werden, d.h. als Korrespondenz zwischen einer erkenntnisunabhängigen Tatsache und einem Satz.

Aus physikalistischer Sicht besteht die wissensunabhängige Welt nur aus physikalischen Entitäten. Es gibt (im Prinzip) eine perfekte wahre Theorie, die alle physikalischen Ereignisse und ihre Relationen beschreibt. Wird die Kausalrelation realistisch interpretiert, dann beschreibt die ideale Theorie auch alle Kausalrelationen in der Welt. Es ist offensichtlich, daß diese ideale Theorie die vollständige Kausalgeschichte jedes beliebigen Ereignisses e in der Welt angibt. Man kann nicht leicht sehen, wie es in einem nicht-trivialen Sinn mehrere solcher vollständigen Kausalerklärungen geben könnte, so daß die Möglichkeit mehrerer *unabhängiger* Erklärungen desselben Ereignisses e eröffnet würde. Dies ist die metaphysische Überzeugung, die hinter dem von Kim herangezogenen Prinzip explanatorischer Exklusion liegt⁶.

Die Exklusivität von Kausalerklärungen

Ich formuliere das Prinzip folgendermaßen:

(PEK) Für kein Ereignis e in der aktuellen Welt w^* gibt es zwei oder mehr Kausalerklärungen, die sowohl vollständig als auch voneinander unabhängig sind.

Aus den metaphysischen Hintergrundüberzeugungen folgt aber, daß die Erklärung eines beliebigen e aus w^* mittels der idealen physikalischen Theorie eine vollständige Erklärung ist. Eine zweite von ihr unabhängige und vollständige Kausalerklärung für e in w^* kann es also nicht geben. Es liegt nun nahe, im Falle psychologischer Erklärungen zu behaupten, hierbei handele es sich nur um unvollständige Erklärungen, die auf die Beschreibung des genauen Kausalgeschehens verzichteten und aus epistemisch-pragmatischen Gründen abstrakte Entitäten (wie z.B. propositionale Ein-

⁶ Vgl. vor allem Kims Aufsatz „Mechanism, purpose, and explanatory exclusion“ in: *J. Kim, Supervenience and Mind. Selected Philosophical Essays*, Cambridge 1993, 237–264.

stellungen) einführen. Nach dieser Auffassung konstruieren wir das mentalistische Vokabular als eine Art abkürzende Beschreibungsweise, weil die Darstellung der zugrundeliegenden physischen Ursachen unsere epistemischen Möglichkeiten überfordern würde. Ein solcher Abstraktionismus droht in eine fikcionalistische Auffassung zu kollabieren. Es drängt sich nämlich die Schlußfolgerung auf, die abstrakten Entitäten seien nur nützliche Fiktionen, die in der geistunabhängigen Welt keine kausale Rolle einnehmen. Damit wäre aber Prinzip [3] und nicht mehr Prinzip [2] negiert. Um diese Möglichkeit auszuschließen, muß die grundlegende Strategie bei der Verneinung von Prinzip [2] darin bestehen, der mentalen Ebene mehr als nur epistemische (folglich metaphysische) Eigenständigkeit zuzugestehen. Kausal relevant können die Beschreibungen im mentalistischen Vokabular nur sein, wenn garantiert wäre, daß sie in einer verlässlichen Beziehung zu den geistunabhängigen Kausalrelationen in der Welt stehen. Die durch die mentalistische Beschreibung eingeführten Entitäten wären dann mehr als bloße Fiktionen.

Eine naheliegende Lösung dieses Problems besteht darin, eine Abhängigkeit der mentalen Ebene von der physischen Ebene anzunehmen, um so die Realität des Mentalen anerkennen zu können, ohne den Primat des Physischen aufzugeben. Das ist grundsätzlich die Strategie für die Negation von Prinzip [2]. Meine These ist, daß alle Versuche, diese Abhängigkeit zu bestimmen, in folgendes Dilemma geraten: *Entweder wird die Abhängigkeit so locker konstruiert, daß die psychologische Kausalerklärung vom realen Kausalgeschehen in der Welt weitestgehend abgelöst wird und die metaphysische Relevanz der Erklärung fragwürdig wird, oder die Abhängigkeit wird so stark, daß die psychologische Kausalerklärung in die physikalische kollabiert.* Im ersten Fall läßt sich ein non-reduktiver Physikalismus aufrechterhalten. Der zweite (und interessantere) Fall ist dann gegeben, wenn die Kausalerklärung auf der psychologischen Ebene in letzter Analyse auf die physikalische Ebene reduziert werden muß. Es muß wiederum klargestellt werden, daß es im vorliegenden Kontext nur um einen metaphysisch interpretierten Reduktionsbegriff gehen kann. Die Tatsache, daß wir aufgrund unserer epistemischen Grenzen die Reduktion eventuell nicht durchführen können, ist aus metaphysisch-realistischer Perspektive nicht ausschlaggebend. Anders ausgedrückt: Ich werde zu zeigen versuchen, daß die metaphysisch gehaltvollen Konzeptionen eines *prima facie* nicht-reduktiven Materialismus in letzter Analyse in ein reduktionistisches Konzept kollabieren. Außerdem werde ich zu zeigen versuchen, daß es innerhalb eines reduktionistischen Ansatzes keinen metaphysisch gehaltvollen Begriff mentaler Verursachung gibt. Bevor ich mich einzelnen Theorien zuwende, will ich die Grundalternative noch etwas klarer herausarbeiten:

(a) *non-reduktive Ansätze*

Verschiedene Autoren haben die Eigenständigkeit des mentalen Bereichs (sowohl des Phänomenbereichs als auch der mentalistischen Begrifflichkeit) herausgearbeitet und auf diesem Hintergrund eine Irreduzibilitätsthese vertreten. Sind z. B. mentale Eigenschaften real und nicht auf physische Eigenschaften reduzierbar, dann läßt sich leichter verdeutlichen, warum mentale Verursachung eine eigenständige Form der Verursachung darstellt. Wenn man annimmt, daß Reales auch kausale Kräfte besitzt (gegen den Epiphänomenalismus), dann folgt, daß reale und irreduzible Entitäten auch reale und irreduzible kausale Kräfte besitzen. Das entscheidende Problem ist dieses: Wenn die kausalen Kräfte der mentalen Entitäten auf irgendeine Weise auf physische Entitäten einwirken, dann ist die Geschlossenheit des physischen Bereiches verletzt, es gibt „downward causation“ von der mentalen auf die physische Ebene. Kim hat überzeugend argumentiert: „Non-reductive physicalism is committed to downward causation.“⁷ Deshalb ist dieser Weg mit der Negation von Prinzip [2] logisch unverträglich.

Non-reduktive Physikalisten haben versucht, diese Konsequenz dadurch zu umgehen, daß sie unter der Hand nur mehr eine epistemische Irreduzibilitätsthese vertraten. Die für Realisten unaufgebbare Unterscheidung zwischen metaphysischer und epistemischer Reduzierbarkeit wurde dabei aber oft genug nicht klar genug gesehen.

(b) *reduktionistische Ansätze*

Schwieriger noch verhält es sich bei der Analyse der verdeckt oder offen reduktionistischen Positionen. Meine These ist, daß sowohl eine starke Supervenienzthese wie auch gängige Formen der Typenidentitätstheorie in letzter Konsequenz reduktionistisch sind. Dadurch ergibt sich ein spezielles Problem. Ist eine zur Diskussion stehende höherstufige Theorie auf eine besser abgesicherte fundamentalere Theorie reduzierbar, dann ist der Realitätsbezug der höherstufigen Theorie in gewisser Weise (wenn auch indirekt) garantiert. Erfolgreiche Reduktion ermöglicht also einen kritischen Realismus in bezug auf die reduzierte Ebene. Auch die mentale Begrifflichkeit bezieht sich also auf etwas, nämlich auf komplexe physische Entitäten. In diesem Sinne kann Reduktion tatsächlich eine Form des mentalen Realismus absichern, ohne die kausale Geschlossenheit des physischen Bereiches zu zerstören. Das Problem liegt hier auf einer sublimeren Ebene. Dadurch, daß mentale Entitäten eigentlich nichts anderes sind als physische Entitäten, verliert die Rede von mentaler Verursachung ihren spezifischen Sinn. Man kann mentale Verursachung nicht dadurch retten, daß man dem mentalen Bereich alle spezifisch mentalen Eigenschaften (z.B. Intentionalität und

⁷ J. Kim, „The Non-Reductivist's Troubles with Mental Causation“, in: *Heil* (Anm. 4), 189–210, hier 203.

Qualia) abspricht. Kim hat dieses Problem mit der für ihn typischen Prägnanz auf den Punkt gebracht.

„... what good is it, one might ask, if you save mental causation but end up losing mentality in the process? (The Vietnam metaphor of saving a village by destroying it comes to mind.)“⁸

Obwohl Reduktion also das probate Mittel ist, um die kausale Geschlossenheit des physischen Bereichs abzusichern, kann die reduktionistische Strategie kein überzeugendes Konzept *mentaler* Verursachung hervorbringen. Diese Thesen sind ohne weitere Analysen vermutlich nicht sofort einsichtig. Im folgenden will ich versuchen, in einer kompakten Form ihre Plausibilität aufzuzeigen. Dazu werde ich exemplarisch einige Vertreter der Grundtypen der Negation von Prinzip [2] herausgreifen. Die fundamentalste Unterscheidung innerhalb der verschiedenen Positionen, die Prinzip [2] negieren, ergibt sich durch die Antwort auf die Frage: Sind die im psychologischen Vokabular beschriebenen Entitäten verschieden von den im physikalischen Vokabular beschriebenen Entitäten? Jede Form der psychophysischen Identitätstheorie antwortet verneinend auf diese Frage. Unter den Theorien, die die Frage positiv beantworten, ist die Supervenienztheorie die derzeit bedeutendste. Zunächst soll gezeigt werden, warum die Supervenienztheorie den psychologischen Kausalerklärungen keine metaphysisch-realistische Interpretation geben kann. Grundlegende Formen der Identitätstheorie (Typenidentität, Funktionalismus und reine Tokenidentität) werden anschließend untersucht.

Supervenienztheorie

Der Kerngedanke der Supervenienztheorie ist der einer Eigenschaftskovarianz: „no difference of one sort without differences of another sort“⁹. Also: keine Veränderung in den supervenienten Eigenschaften ohne Veränderung in den subvenienten Eigenschaften. Diese bloße Kovarianzannahme ist jedoch zu unspezifisch, um die Supervenienztheorie als klare Position hervortreten zu lassen. Gesucht wird eine Beziehung zwischen der super- und der subvenienten Ebene, die stark genug ist, um über bloß zufällige Kovarianz hinaus eine genuine *Abhängigkeit* zu begründen. Ein erster Schritt in diese Richtung besteht darin, nur modal starke Formen von Kovarianz zuzulassen. Was damit gemeint ist, kann durch den Einbezug möglicher Welten in die entsprechenden Definitionen bestimmt werden. Ein nur auf *eine einzige* (die aktuelle) Welt bezogener Kovarianzbegriff würde es zulassen, daß z. B. in einer Zwillingswelt, die der aktuellen ansonsten detailgetreu gleicht, die psychologischen Eigenschaften ganz anders verteilt wären oder aber überhaupt nicht existierten (schwache Supervenienz). Darüber hinaus muß die Kovarianzbeziehung hinreichend spezifisch sein, um eine echte Abhängigkeit be-

⁸ Kim, Supervenience 367.

⁹ Siehe vor allem Kim (ebd.), für klassische Aufsätze zur Supervenienztheorie.

gründen zu können. Ein Kovarianzbegriff, der sich immer nur global auf ganze Welten bezieht, könnte nicht verhindern, daß sich in *einer Welt* zwei bis auf ihre Raum-Zeit-Stellen physisch ununterscheidbare Individuen in ihren psychologischen Eigenschaften drastisch unterscheiden (globale Supervenienz). Ein für das Leib-Seele-Problem verwendbarer Kovarianzbegriff sollte daher sowohl weltübergreifend als auch auf individuelle Eigenschaften bezogen konzipiert werden (starke Supervenienz). Ohne technische Details kann man den dazu benötigten Kovarianzbegriff intuitiv darstellen.

Nehmen wir an, es gäbe eine Menge von supervenienten (z. B. mentalen) Eigenschaften und eine Menge von subvenienten (z. B. physischen) Eigenschaften. Im Falle starker Kovarianz dieser Eigenschaften muß für beliebige Entitäten und beliebige Welten gelten:

- (SK) Für beliebige Welten w_i und w_j und beliebige Entitäten x und y :
 Wenn x in w_i bezüglich seiner subvenienten Eigenschaften ununterscheidbar ist von y in derselben Welt w_i , oder einer anderen Welt w_j , dann ist x in w_i auch in seinen supervenienten Eigenschaften ununterscheidbar von y in w_i oder w_j .

Die reine Kovarianz von Eigenschaften könnte zunächst ganz symmetrisch konzipiert werden. Der Supervenienzbegriff enthält aber eine Prioritätsannahme bezüglich der subvenienten Ebene. In die Formulierung von (SK) ging diese Annahme schon implizit ein. (SK) behauptet, daß aus der Ununterscheidbarkeit der subvenienten Eigenschaften die Ununterscheidbarkeit der supervenienten Eigenschaften folgt. Damit soll die metaphysische Priorität der physikalischen Ebene über die psychologische Ebene ausgedrückt werden. Die umgekehrte Beziehung gilt nicht. Entitäten, die in ihren supervenienten Eigenschaften ununterscheidbar sind, können durchaus in ihren subvenienten Eigenschaften unterscheidbar sein (multiple Realisierung). Erst wenn diese Asymmetrie explizit in die Definition der Supervenienzrelation eingegangen ist, hat man neben bloßer Eigenschaftskovarianz auch eine echte Abhängigkeitsbeziehung eingeführt. Eine solche aus metaphysischen Gründen angenommene Asymmetrie oder Abhängigkeit bedarf der Explikation. In welchem Sinne hat die physikalische Ebene einen Primat über die psychologische Ebene? Der beste Weg scheint über den Einbezug mereologischer Überlegungen zu verlaufen. Superveniente Makroeigenschaften werden durch subveniente Mikroeigenschaften determiniert. Auf der Mikroebene identische Welten sind notwendig auf der Makroebene identisch. In umgekehrter Richtung liegt keine Determination vor.

Diese Kombination von Kovarianz, Abhängigkeit und multipler Realisierbarkeit sieht vielversprechend aus. Die superveniente Ebene genießt aufgrund ihrer Irreduzibilität eine gewisse Autonomie. In diesem Sinne könnte sie auch eine von der subvenienten Ebene verschiedene kausale Rolle einnehmen. Der Gedankengang könnte so verlaufen: Jeder mentalen Verände-

rung entspricht eine physikalische Veränderung. Jeder physikalischen Veränderung entspricht eine Änderung in den kausalen Beziehungen. Also sind die mentalen Veränderungen relevant für die kausalen Beziehungen.

Dieser Gedankengang ist aber irreführend. Er läßt es so aussehen, als ob die Abhängigkeit von der supervenienten zur subvenienten Ebene verlief. Daß jeder mentalen Veränderung eine physische Veränderung entspricht (Kovarianz im allgemeinsten Sinne), impliziert nicht, daß physische Eigenschaften von mentalen Eigenschaften abhängig sind. Die Abhängigkeitsbeziehung verläuft gerade umgekehrt. Sie ist (im oben entwickelten Sinne) sogar stark genug, um in einem spezifischen Sinne eine Reduktion der supervenienten Eigenschaften auf die subvenienten Eigenschaften zu garantieren. Da die mentalen Eigenschaften multipel realisierbar sind, könnten sie – einem Vorschlag Kims entsprechend¹⁰ – auf eine Disjunktion mehrerer physischer Eigenschaften reduziert werden. Solche Disjunktionen sind zugegebenerweise etwas eigenartige Gebilde. Vor allem läßt sich mit Recht fragen, ob wir jemals in der Lage sein werden, die vollständige Disjunktion anzugeben, die man für die Angabe der Brückengesetze beim Reduktionsverfahren bräuchte. Aber das ist wieder eine epistemische Fragestellung. Metaphysisch gesehen, wäre gegen die Vorstellung einer strengen nomologischen Korrelation zwischen einer Eigenschaft und einer Disjunktion von anderen Eigenschaften nichts einzuwenden. Es wurde auch angezweifelt, ob die Koextensionalität des mentalen Prädikats mit der Disjunktion physikalischer Prädikate überhaupt hinreichend sei, um Brückengesetze zwischen den beiden Theorieebenen zu formulieren¹¹. Es wird dann fraglich, ob Reduzierbarkeit im klassischen Sinne (über Brückengesetze) überhaupt gewährleistet ist. Ich glaube allerdings, daß die Intuition Kims von solchen Schwierigkeiten bezüglich ihrer technischen Explizierbarkeit nicht im Kern entkräftet wird. Es geht ihm darum, daß das mentale Prädikat sich auf nichts anderes bezieht als die Disjunktion der physikalischen Prädikate. Das mentale Prädikat beschreibt „nichts anderes als“ die Disjunktion der physischen Prädikate. Genau das ist der intuitive Begriff von Reduzierbarkeit.

An Kims Argument läßt sich auch eine Grundschwierigkeit aller Ansätze, die multiple Realisierung erlauben, darstellen: Jedem Einzelfall der Instantiierung einer mentalen Eigenschaft liegt nicht die ganze Disjunktion, sondern immer nur *eine* physikalische Eigenschaft zugrunde. Durch sie allein ist auch das Kausalgeschehen in diesem Einzelfalle bestimmt. Daß sich die mentale Eigenschaft nur auf eine Disjunktion mehrerer physischer Eigenschaften reduzieren läßt, zeigt deutlich, daß sie vom wirklichen Kausalgeschehen abstrahiert. Daß es für die höherstufige Eigenschaft unwichtig ist,

¹⁰ Vgl. dazu Kims Artikel „Multiple Realization and the Metaphysics of Reduction“, in: *Kim, Supervenience* 309–35.

¹¹ Eventuell sind solche Disjunktionen keine gesetzesfähigen Prädikate, weil sie beispielsweise nicht (im Sinne Goodmans) projizierbar sind.

wie sie genau realisiert ist, belegt nicht die kausale Irrelevanz der Realisierung, sondern im Gegenteil die Entfernung der höherstufigen Eigenschaft vom Kausalgeschehen.

Es läßt sich noch auf eine andere Weise die These untermauern, daß starke Supervenienz tatsächlich eine metaphysische Reduktionsthese impliziert. In letzter Konsequenz (aus der Gottesperspektive) behauptet die Supervenienztheorie, daß nur das Physische wirklich existiert, selbst wenn wir (aufgrund unserer epistemischen Begrenzungen) mentale Eigenschaften niemals auf physische Eigenschaften reduzieren können. Diese These kann durch eine semantisch orientierte Überlegung erhärtet werden: Betrachten wir die Menge der psychologischen Aussagen als eine Theorie. Diese Theorie hat Realisationen oder Modelle, die für jede Aussage einen Wahrheitswert festlegen. Dasselbe gilt für die Physik. Wenn die Psychologie der Physik gegenüber supervenient sein soll, dann muß sichergestellt sein, daß Modelle der physikalischen Theorie jeder Aussage der Psychologie einen Wahrheitswert zuordnen (aber nicht umgekehrt). Wenn die Wahrheitswerte der psychologischen Aussagen nicht letztendlich von den Realisationen der physikalischen Theorie determiniert werden, hat der asymmetrische Charakter der Supervenienzrelation keine ontologische Grundlage. Die wahren Aussagen der psychologischen Theorie müssen daher in der physikalischen Theorie Entsprechungen haben. Anders ausgedrückt: Durch die Bestimmung der physikalischen Fakten werden alle psychologischen Fakten bestimmt. Modelle der physikalischen Theorie sind also auch Modelle der psychologischen Theorie. Die Frage ist nun, wie zwei Theorien mit derart verschiedenem Vokabular so aufeinander bezogen werden können, daß die Modelle der einen auch die Modelle der anderen sein können. Auf der begrifflichen Ebene scheinen die beiden Theorien beziehungslos nebeneinander zu stehen. Wenn aber auch ihre ontologischen Realisationen beziehungslos nebeneinander stehen, kollabiert die Supervenienzthese. Bonevac hat darauf hingewiesen, daß man eine dritte (eine Supertheorie) bräuchte, mit deren Hilfe man erklären kann, daß eine der beiden Theorien tatsächlich vollkommen in der Ontologie der anderen Theorie aufgeht. Würden die psychologische Theorie und die physikalische Theorie beide in die Supertheorie übersetzt, dann ließe sich zeigen, daß die Menge der von der psychologischen Theorie eingeführten Entitäten eine Teilmenge der von der physikalischen Theorie eingeführten Entitäten darstellt. Genau dies ist nach seiner Auffassung die Intuition einer starken Supervenienzthese.

„This not only explicates talk of determination and indiscernibility across possible worlds but makes clear a sense in which those notions are too weak. After all, the advocate of the ontological significance of supervenience wants to claim not merely that the physical determines the mental, say, but that the physical is all there is; the mental, in the end, is just the physical.“¹²

¹² D. Bonevac, Supervenience and Ontology, in: APQ 25 (1988) 37–47, hier 46.

Auf diesem Hintergrund wird dann auch klar, warum eine starke Supervenienzthese notwendig in einem metaphysischen Sinne reduktionistisch ist. Die Tatsache, daß wir die psychologische Theorie nicht auf die physikalische Theorie reduzieren können, beruht auf einer rein epistemischen Beschränkung. In letzter Analyse (aus der Gottesperspektive) werden die Wahrheitswerte aller Aussagen in der psychologischen Theorie durch die Realisationen der physikalischen Theorie bestimmt, weil die psychologische Theorie im Prinzip in die physikalische Theorie übersetzbar ist. Bonevac schließt mit Recht daraus, Supervenienz sei lediglich „reducibility in the long run“¹³. Selbst wenn also die Autonomie der psychologischen Erklärungen durch faktische Irreduzibilität gesichert wäre, so könnte man doch den mentalen Eigenschaften als solchen aus einer metaphysischen Perspektive keine kausale Relevanz zusprechen.

Der – schon erwähnte – wichtigste Einwand gegen eine Strategie dieser Art würde darauf hinweisen, daß es gerade die Reduzierbarkeit ist, welche die kausale Relevanz garantiert. Wenn das psychologische Vokabular nämlich reduzierbar ist, dann bezieht es sich erfolgreich auf die Welt. Die psychologische Theorie muß daher nicht als falsch verworfen werden (Elimination). Die entscheidende Frage ist aber, ob ein solcher durch Reduzierbarkeit garantierter „mentaler“ Realismus den mentalen Eigenschaften *qua mentalen* Eigenschaften eine kausale Rolle zusprechen kann. Das kann mit guten Gründen bezweifelt werden. Es scheint vielmehr so zu sein, daß die mentalen Eigenschaften nur *qua* ihrer Reduzierbarkeit eine kausale Rolle haben. Strenggenommen gibt es dann keine mentale Verursachung als solche. Ausnahmslos alle Kausalbeziehungen werden von einer idealisierten Theorie der physischen Welt lückenlos beschrieben. Da wir über diese Theorie nicht verfügen, ist die Einführung mentaler Entitäten für uns explanatorisch von Nutzen. Über die prinzipiell mögliche Reduktion ist der psychologischen Sprache also der Realitätsbezug gesichert. Was nicht gesichert ist, ist eine gegenüber der physischen Welt in irgendeinem metaphysisch relevanten Sinne kausale Autonomie der mentalen Ebene. Nach meiner Auffassung setzt ein gehaltvoller Begriff mentaler Verursachung die metaphysische Irreduzibilität der psychologischen Ebene voraus. Genau dies aber kann ein starker Supervenienzbegriff nicht garantieren.

Typenidentitätstheorie

Psychologische Identitätstheorien scheinen das Problem mentaler Verursachung auf effektive Weise lösen zu können: Identische Entitäten können miteinander nicht um die kausale Rolle konkurrieren. Betrachtet man die aktuell diskutierten Identitätstheorien genauer, wird es auch hier zweifelhaft, ob in ihnen dem Mentalen *qua Mentalem* eine kausale Rolle zukommt.

¹³ Ebd.

Auch diese These kann ich nur exemplarisch erhärten. Beginnen will ich mit der Typenidentitätstheorie. Die aktuell einflußreichste Form der Typenidentitätstheorie wurde von Armstrong und Lewis entwickelt¹⁴. Ich werde die Grundintuitionen dieser Theorie als Gegenstand der Analyse wählen.

Ein Charakteristikum dieser Theorieform liegt darin, daß sie eng mit dem wissenschaftlichen Realismus verbunden ist und sich gegen jede apriorische Lehnstuhlphilosophie wendet. Die Alltagspsychologie wird als eine primitive protowissenschaftliche Theorie begriffen. In der Alltagspsychologie greifen wir über ihre kausalen Rollen bestimmte mentale Eigenschaften heraus. Die physikalistische Grundannahme (kausale Geschlossenheit) fordert nun, daß alle kausalen Rollen durch rein physikalische Eigenschaften übernommen werden. Das Problem wird gelöst durch die Annahme, daß mentale Eigenschaften identisch sind mit physischen Eigenschaften, die dieselbe kausale Rolle einnehmen. Die fragliche kausale Rolle fällt also nicht zwei verschiedenen Entitäten gleichzeitig zu, was einen Konflikt mit dem kausalen Exklusionsprinzip heraufbeschwören würde. In der Alltagspsychologie greifen wir also eine bestimmte Eigenschaft mittels ihrer kausalen Rolle heraus und identifizieren sie dann mit einer empirisch entdeckten physischen Eigenschaft, die dieselbe kausale Rolle einnimmt. Die Identifizierung von mentalen mit physischen Zuständen entspricht strukturell der Identifizierung von Genen mit DNA oder von Wasser mit H₂O. Diese Beispiele zeigen bereits deutlich, daß die beiden Relata nicht in jeder Hinsicht den gleichen Status haben. Nur in der physikalischen Sprache lassen sich die kausalen Zusammenhänge kleinteilig aufklären. Der *modus operandi* der kausalen Wechselwirkung kann nur auf der physischen Ebene genau dargestellt werden. Die psychologische Terminologie scheint für sich genommen völlig ungeeignet, um die kausalen Zusammenhänge im Detail zu beschreiben. Die mentalen Prädikate sind in diesem Sinne nicht erweiterbar (extendible). Eine ziemlich grobe Erklärung in der Physik kann durch immer kleinteiligere, präzisere Beschreibung erweitert werden, bis man schließlich die fundamentalen Entitäten, Größen und Gesetze beschreibt. Selbst wenn wir im Alltag Wasser sehr gut über seine kausalen Rollen identifizieren können, so erwarten wir doch die Antwort auf die Frage, was Wasser eigentlich ist, durch eine Beschreibung auf der molekularen Ebene.

Ein kurzer Blick auf Lewis' Theorie macht diese Asymmetrie noch deutlicher: Er betrachtet die Alltagspsychologie als eine termeinführende Theorie. Diese Theorie sagt uns (unter anderem) etwas über die kausalen Beziehungen zwischen mentalen Zuständen, verschiedenen sensorischen Eingä-

¹⁴ Für eine kurze Darstellung von *Armstrongs* Theorie siehe z. B. seinen Aufsatz „The Causal Theory of the Mind“, in: NHP 11 (1977) 82–95. Die wichtigsten einschlägigen Aufsätze von David Lewis finden sich zusammengestellt in *D. Lewis, Philosophical Papers, Volume I*, Oxford 1983, besonders in den Teilen „Ontology“ und „Philosophy of Mind“ (1–160).

ben (input) und beobachtbarem Verhalten (output). Das funktionalistische Element dieses Vorschlages ist unübersehbar. Eine zweite Theorie (die physische oder neurophysiologische) realisiert oder erfüllt die erste Theorie. Entscheidend für die Durchführung dieser Korrelation ist, daß die erste Theorie von ihrem psychologischen Vokabular befreit wird (topic neutrality). Zu diesem Zweck ersetzt man alle psychologischen Terme durch gebundene Variable. Damit wird behauptet, daß es eine geordnete Menge von Entitäten gibt, welche die erste Theorie erfüllt. Diese Erfüllung leistet dann auf eindeutige Weise die geordnete Menge physikalischer Eigenschaften oder Zustände, die durch die zweite (physikalische) Theorie gegeben wird. Dieses Verfahren ermöglicht die Reduktion der Alltagspsychologie auf die Neurophysiologie, ohne daß die Brückengesetze noch auf mentalistisches Vokabular Bezug nehmen. Lewis selbst folgert, daß in solchen Fällen Reduktion nicht nur möglich, sondern unvermeidlich sei¹⁵. Asymmetrische Reduzierbarkeit schließt aber eine in irgendeinem Sinne eigenständige kausale Rolle der reduzierten Ebene aus. Gerettet wird wieder nur eine Ausdrucksweise, nicht ein metaphysisch relevantes Faktum. Eine weitere Überlegung kann diese These zusätzlich untermauern.

Bekanntlich ist es schwer, eine Typenidentitätstheorie von der Art, wie Lewis sie vertritt, genau zu charakterisieren, da sie Elemente des Funktionalismus bewußt in sich aufnimmt. Verschiedene Autoren haben gezeigt, daß diese Amalgamierung zu logischen Unverträglichkeiten führt¹⁶. Die zentrale Frage ist, ob die kausalen Rollen nur dazu dienen, identische Eigenschaften in einer gegebenen Welt herauszufinden (Identifikation) oder ob die kausalen Rollen selbst es sind, die die Identität von Eigenschaften bestimmen. Aus funktionalistischer Sicht ist eine mentale Eigenschaft essentiell durch ihre kausale Rolle bestimmt und nur kontingent mit ihrer physischen Realisierung verbunden. In einer reinen Typenidentitätstheorie hingegen kann die kausale Rolle allenfalls die Referenz eines mentalen Typs festlegen, nicht aber seine Essenz beschreiben¹⁷. Da ich weiter unten noch auf den Funktionalismus eingehen werde, beschränke ich mich nur auf den typentheoretischen Aspekt in einer Theorie von Armstrong-Lewis-Typ. Hier geschieht der Rekurs auf funktionalistische Intuitionen unter epistemischer Rücksicht: Die funktionale Rolle dient dazu, eine mentale Eigenschaft *herauszugreifen*. Ihre Essenz wird über die Typenidentität mit einer physischen Eigenschaft bestimmt.

Nehmen wir zur Exemplifizierung des Gedankenganges das Kunstwort

¹⁵ D. Lewis, „How to define theoretical terms“, in: Lewis (Anm. 14), 78–95, hier 78.

¹⁶ Vgl. z.B.: C. McDonald, *Mind-Body Identity Theories*, London/New York 1989, 80 f.: „Evidently, however, commitment to type-type identities is ultimately incompatible with the claim that the nature of any given mental state is causal-functional. ... the causal role definitive of a given mental type may fix the *reference* of that type, but it does not provide its essence.“

¹⁷ Unter den *termini technici* „Referenz“ und „referieren“ verstehe ich die Bezugnahme von Zeichen auf Bezeichnetes. Dieser Anglizismus hat sich in der analytischen Philosophie eingebürgert.

„Dormitivität“. Für das Verständnis dieses Wortes können zwei mögliche Lesarten angegeben werden¹⁸:

Dormitivität₁ = die Eigenschaft, die Schlaf erzeugt

Dormitivität₂ = eine Eigenschaft besitzen, die Schlaf erzeugt.

Lewis konstruiert funktionale Prädikate wie „Dormitivität₁“, d.h. als Prädikate erster Ordnung. Nach seiner Auffassung sollen die entsprechenden Prädikate (wenigstens innerhalb eines fest umgrenzten Rahmens) eindeutig referieren. Setzen wir beispielsweise einen solchen Rahmen und sprechen nur über die Dormitivität₁ eines benzodiazepinhaltigen Medikaments (z.B. Valium) und nennen sie „Dormitivität_{1bzd}“. Nehmen wir an, eine vollständige (ideale) Analyse der Wirkungen der Benzodiazepine gibt uns eine eindeutige Bestimmung dessen, worauf „Dormitivität_{1bzd}“ referiert: die chemische Reaktion XYZ.

Es ist nun offensichtlich, daß das, was „Dormitivität_{1bzd}“ herausgreift, keine multipel realisierbare funktionale Eigenschaft ist, sondern unmittelbar eine physische Eigenschaft. Es ist deshalb in einem gewissen Sinne völlig unproblematisch zu sagen: „Dormitivität_{1bzd} verursacht Schlaf“. Das bedeutet aber nicht, daß die funktionale Eigenschaft „Dormitivität_{1bzd}“ als solche eine in der Welt kausal wirksame Eigenschaft wäre. Es ist die chemische Reaktion XYZ, die Schlaf erzeugt. Die funktionale Beschreibung hatte nur epistemische Bedeutung (beim Herausgreifen), die Aufdeckung des durch Forschung kleinteilig erweiterbaren Kausalzusammenhanges liefert allein die Analyse von XYZ. Dormitivität_{1bzd} kann und muß auf XYZ reduziert werden.

Wenn die funktionale Beschreibung unmittelbar und eindeutig auf nicht-funktionale Zustände referiert, kommt der funktionalen Ebene als solcher keine eigenständige kausale Relevanz mehr zu. Im vorliegenden Kontext geht es darum, über den Aufstieg auf die funktionale Ebene dem Mentalen eine kausale Rolle zugestehen zu können, d.h. den mentalen Zuständen kausale Relevanz zubilligen zu können. Das Ganze wirkt fast wie ein Illusionstrick. Weil uns die psychologische Begrifflichkeit als eine Redeweise erhalten bleibt, glauben wir, der mentalen Ebene käme auch kausale Relevanz zu. Im Hintergrund steht aber eine Typenidentitätstheorie mit einem reduktionistischen Anliegen, d.h. der Überzeugung, daß in letzter Analyse nur das Physische existiert. Auch Block weist darauf hin, daß Lewis nichts mehr rettet als eine nützliche Sprechweise. Das eigentliche Kausalgeschehen läuft ausschließlich auf der physischen Ebene ab.

„You do not vindicate the causal efficacy of the properties of a special science by construing its terms as referring to the properties of other sciences. The Lewis proposal does preserve a way of talking that we want.“¹⁹

¹⁸ Vgl. für dieses Beispiel auch: N. Block, „Can the Mind Change the World?“, in: George Boonlos (ed.): *Meaning and Method. Essays in the Honor of Hilary Putnam*, Cambridge 1990, 163 f.

¹⁹ Ebd. 165.

Mit diesen Argumenten ist auch schon der wesentlichste Einwand gegen eine rein funktionalistische Theorie mentaler Verursachung angesprochen, insofern sie die funktionale Ebene über den Begriff der kausalen Rolle bestimmt.

Funktionalismus

Der Funktionalismus bietet über die Identifizierung der mentalen mit der funktionalen Ebene die Chance einer Autonomie des Mentalen gegenüber dem Physischen gerade in dem Maße, wie sich die funktionale Ebene ihrer Realisierung gegenüber als autonom erweist²⁰. Für die Entstehung des Funktionalismus halte ich folgende Intuition für zentral: Es gibt eine Typenidentität zwischen mentalen und funktionalen Eigenschaften, zwischen funktionalen und physischen Eigenschaften hingegen nicht. Die Instantiierung einer funktionalen Eigenschaft geschieht nur in einem konkreten physischen Vorkommnis (Tokenidentität). Ohne auf die Geschichte des Funktionalismus einzugehen, will ich mich sofort der momentan einflußreichsten Variante, dem Kausale-Rolle-Funktionalismus zuwenden. Der Grundgedanke ist, daß die Identität funktionaler Eigenschaften nicht wie in anderen funktionalistischen Systemen über Zustände einer Turingmaschine, sondern über ihre kausale Rolle bestimmt wird. Der z.B. von Searle vorgebrachte Einwand, daß algorithmisch spezifizierte Computerzustände überhaupt keine intrinsischen Zustände eines Systems sind, sondern vielmehr von einer Interpretation von außen abhängen, trifft dann nicht mehr. Wenn man die funktionalen Rollen an den Zement des Universums (Kausalität) anbindet, dann scheint ihre geistunabhängige Realität gesichert zu sein.

Es ergibt sich aber sofort das Problem, wie man garantieren kann, daß die Beschreibungen in der Alltagspsychologie sicher und eindeutig die kausalen Rollen herausgreifen, die man braucht, um die funktionalen Eigenschaften zu bestimmen. Gelänge dies nicht, so könne man den Funktionalismus nicht mehr vom Behaviorismus unterscheiden, den er doch überwinden wollte. Es wäre nämlich völlig irrelevant, wie ein System intern aufgebaut ist, solange es sich nur so verhält, daß es die Alltagspsychologie erfüllt (d.h. auf bestimmte Inputs mit bestimmten Outputs reagiert). Die Behauptung, daß etwas diese oder jene funktionale Struktur hat, wäre äquivalent mit der Behauptung, daß es sich so verhält, als hätte es diese Struktur. Aber sehen wir von diesem schwierigen Problem einmal ab und nehmen die funktionale Beschreibungsebene als problemlos gegeben an.

Es fragt sich dann sofort wieder, ob man die Zuschreibung funktional spezifizierter kausaler Rollen wirklich so unmittelbar realistisch interpretieren kann. Die Frage, auf was sich die funktionale Beschreibung eigentlich in der Welt bezieht, wird nicht leicht zu beantworten sein. Es scheint mir aus

²⁰ Siehe *H. Putnam, Mind Language and Reality. Philosophical Papers. Vol. II* (Cambridge, Mass. 1975), für einige klassische Texte des Funktionalismus.

der dem Funktionalismus inhärenten Idee der multiplen Realisierbarkeit ableitbar zu sein, daß die funktionale Beschreibung gerade von *de facto* ablaufenden Kausalgeschehen abstrahieren muß. Die funktionale Beschreibung findet auf einer sehr abstrakten Ebene statt. Man sollte sehr vorsichtig sein, sie direkt realistisch zu interpretieren. Eine sehr generelle Frage verdeutlicht diesen Punkt: Wie kann eine Eigenschaft, die *ex hypothesi* eine kausale Rolle ist, ihrerseits noch eine Ursache sein? Oder: Wie kann die kausale Rolle selbst kausal wirksam sein? Kausale Rollen befinden sich im Verhältnis zu Ursachen auf einer Metaebene. Es ist daher äußerst fraglich, ob funktionale Eigenschaften als solche kausal wirksam sein können.

Funktionalisten haben funktionalen Eigenschaften jedoch kausale Wirksamkeit zugesprochen. Putnam ging (als er noch Funktionalist war) sogar so weit, zu behaupten:

„... the ultimate constituents don't matter ... only the higher level structure matters ... The same explanation will go in any world (whatever the microstructure) in which those higher level structural features are present.“²¹

Der entscheidende Punkt liegt darin, daß er hier von Erklärungen spricht. In einem metaphysisch-realistischen Ansatz muß man aber deutlich zwischen Kausalerklärungen (epistemische Ebene) und Kausalrelationen (metaphysische Ebene) unterscheiden. Daß wir für erfolgreiche Kausalerklärungen meist die genauen Kausalrelationen nicht kennen, bzw. sogar von ihnen abstrahieren müssen, ist offenkundig. In diesem Sinne nimmt die Beschreibung auf der funktionalen Ebene keine Rücksicht auf das genaue Kausalgeschehen. Es spielt im epistemischen Zusammenhang keine Rolle, welche Mikroeigenschaften die funktionale Beschreibung realisieren. Das Kausalgeschehen läuft hingegen nur innerhalb der durch Tokenidentität mit der funktionalen Ebene verbundenen konkreten Realisierung (Instantiierung) ab. Obwohl eine künstliche Niere und eine natürliche auf funktionaler Ebene äußerst ähnlich sind, ist das wirkliche Kausalgeschehen denkbar verschieden. Zwischen der abstrakten funktionalen Ebene und der konkreten Realisierung tut sich ein ontologischer Bruch auf. In diesem Sinne rechtfertigt auch der Funktionalismus nur einen „way of talking“ aus einer bestimmten instrumentell-pragmatisch relevanten Perspektive („from a stance“). Er garantiert eine irreduzible epistemische Ebene der Kausalerklärung, vermittelt uns aber keine neuen Einsichten über den Zement des Universums. Schematisch könnte man den Haupteinwand so darstellen:

(1) Funktionale Beschreibungen individuieren Zustände, die eine Vielzahl von Realisationsmöglichkeiten abdecken.

(2) Wenn ein Zustand eine Vielzahl von Realisationsmöglichkeiten abdeckt, dann ist nicht er es, sondern die einzelnen Realisationen, die kausal relevant sind.

²¹ Ebd. 296.

(3) Also individuieren funktionale Beschreibungen keine kausal relevanten Zustände.

Daher sollte man funktionale Beschreibungen nicht unmittelbar realistisch interpretieren. Dennett hat wiederholt argumentiert, daß den funktionalen Zuschreibungen objektiv in der Natur nichts entspreche. Sie seien intrinsisch an einen explanatorischen Standpunkt gebunden²².

Es könnte eingewendet werden, der Funktionalismus garantiere die kausale Relevanz funktional beschriebener mentaler Eigenschaften trotzdem über die vorausgesetzte Tokenidentitätstheorie. Es ist aber wiederum fraglich, wie die Identitätsthese unter den gemachten Voraussetzungen (kausale Geschlossenheit und Exklusivität) den mentalen Eigenschaften kausale Kraft garantieren kann. Die folgenden Annahmen scheinen zumindest *prima facie* eine kausale Rolle mentaler Eigenschaften auszuschließen:

(1) *Tokenidentität*: Jedes Ereignis mit mentalen Eigenschaften ist identisch mit einem Ereignis mit physischen Eigenschaften.

(2) *Kausale Exklusivität*: Es gibt für ein Ereignis genau eine vollständige kausale Vorgeschichte. Es gibt keine Überdetermination.

(3) *Kausale Abgeschlossenheit*: Die kausalen Kräfte eines Ereignisses sind vollständig determiniert durch seine physischen Eigenschaften.

Aber diese Problematik läßt sich am einfachsten bei einer reinen Tokenidentitätstheorie (ohne Funktionalismus) analysieren.

Tokenidentitätstheorie

Die Tokenidentitätstheorie soll am Beispiel von Donald Davidson nochmals ganz unabhängig vom Funktionalismus diskutiert werden. Er hat mit seinem Lösungsvorschlag für das Problem mentaler Verursachung die aktuelle Debatte geprägt wie kaum ein anderer²³. Da in Davidsons Theorie jede Kausalrelation zwischen zwei Ereignissen unter ein Kausalgesetz subsumiert sein muß, droht die anomale mentale Ebene letztendlich kausal irrelevant zu bleiben. Der Ausweg liegt in einer metaphysischen Tokenidentitätstheorie: die in psychologischer Begrifflichkeit beschriebenen Ereignisse sind identisch mit den in physikalischer Begrifflichkeit beschriebenen Ereignissen und können deshalb nicht mit ihnen um die kausale Wirkmächtigkeit konkurrieren. Seine Tokenidentitätstheorie wurde von vielen Autoren erfolgreich kritisiert, weil es äußerst fraglich ist, ob sie den mentalen Eigenschaften eine kausale Rolle zumessen kann. Der Epiphänomenalismusvorwurf gegenüber Davidson wurde durch Argumente der folgenden Art gestützt²⁴:

²² „Mother Nature doesn't commit herself explicitly and objectively to any functional attributions; all such attributions depend on the mind-set of the intentional stance.“ *D. Dennett, The Intentional Stance*, Cambridge, MA 1987, 320.

²³ Die wichtigsten Texte finden sich in: *D. Davidson, Essays on Actions and Events*, Oxford 1980.

²⁴ Vgl. dazu: *R. van Gulick*, „Who's in Charge Here? And Who's Doing all the Work?“, in: *Heil* (Anm. 4) 233–258, hier 235 ff.

Ableitung I:

- (1) *Nomologischer Charakter der Kausalität*: Ein Ereignis e_1 mit der Eigenschaft F kann ein Ereignis e_2 nur dann verursachen, wenn e_2 eine Eigenschaft G besitzt und es ein striktes Gesetz gibt, unter das die beiden Ereignisse aufgrund F und G subsumiert werden können.
- (2) *Anomalie des Mentalen*: Mentale Eigenschaften gehen nicht in strikte Gesetze ein.
- (3) Also: Dadurch, daß e_1 die mentale Eigenschaft M besitzt, kann kein Ereignis e_2 bewirkt werden. Dadurch daß e_1 eine mentale Eigenschaft M besitzt, kann nicht bewirkt werden, daß ein Ereignis e_2 eine Eigenschaft G besitzt. Somit sind mentale Eigenschaften kausal wirkungslos.

Ableitung II:

- (1) *Tokenidentität*: Jedes Ereignis (token) mit mentalen Eigenschaften ist identisch mit einem Ereignis (token) mit physischen Eigenschaften.
- (2) *Kausale Abgeschlossenheit*: Die kausalen Kräfte eines Ereignisses sind vollständig determiniert durch seine physischen Eigenschaften.
- (3) *Nichtreduzierbarkeit des Mentalen*: Mentale Eigenschaften sind nicht reduzierbar auf physische Eigenschaften.
- (4) Also: Mentale Eigenschaften sind kausal wirkungslos.

Die Versuche, Davidsons Ansatz zu verteidigen, konzentrieren sich konsequenterweise darauf, unter Heranziehung von Zusatzannahmen (z. B. Supervenienz) zu belegen, daß die mentalen Eigenschaften auch in einer Theorie Davidsonschen Typs nicht wirkungslos bleiben müssen. Meines Erachtens ist diese Debatte um Eigenschaften im Rahmen einer reinen Tokenidentitätstheorie aber nicht der wesentlichste Teil der Auseinandersetzung mit einer Theorie dieses Typs. Davidson könnte diesen Streitigkeiten um die kausale Kraft von *Eigenschaften* mit stoischer Gelassenheit zuschauen. Sein Programm heißt: Ontologische Reduktion ohne begriffliche Reduktion. Die für eine metaphysisch-realistische Position zentrale Unterscheidung von begrifflicher und ontologischer Ebene spielt für ihn eine oft nicht genügend gewürdigte Schlüsselrolle.

Die vielleicht fundamentalste Grundintuition bei der Lösung des Problems mentaler Verursachung ist für Davidson die Unterscheidung zwischen der rein extensionalen, beschreibungsunabhängigen Gegebenheit partikulärer Ereignisse einerseits und den begrifflichen Mitteln, mit denen wir Ereignisse charakterisieren und herausgreifen, andererseits. Davidson führt diese Unterscheidung mit solcher Stringenz durch, daß die Ereignisse „an sich“ völlig beschreibungstranszendent bleiben. Keine Beschreibung beantwortet letztlich die Frage: „Was ist diese Entität?“ Die Beschreibung kann nur dazu dienen, die Entität herauszugreifen. Die Konkurrenz zwischen psychologischen und physikalischen Beschreibungen betrifft also letztendlich nicht die metaphysische Ebene. Beide Beschreibungen können dasselbe

Ereignis herausgreifen. Das beschreibungstranszendente Ereignis ist gegenüber prädikativen Zuschreibungen wie „mental“ oder „physisch“ völlig neutral. Dementsprechend stellt sich auch das Problem der mentalen Verursachung nur auf der begrifflichen Ebene. Auf der metaphysischen Ebene gibt es nur ein einziges Ereignis, das durch zwei aufeinander irreduzible Beschreibungen herausgegriffen wurde. Setzt man die Realität mit unseren begrifflichen Schemen gleich, so stattet man die Realität mit der unüberbrückbaren Dichotomie dieser verschiedenen begrifflichen Schemen aus. Man könnte auch plakativ sagen, daß in einer so verstandenen Tokenidentitätstheorie die idealistische Nicht-Differenzierung von epistemischer und metaphysischer Ebene die eigentliche Wurzel des Leib-Seele-Problems ist. Es ist offensichtlich, daß ein solcher Ansatz den mentalen Eigenschaften (qua Eigenschaften) keine kausale Kraft zubilligen kann.

Aber dieser Einwand trifft den Ansatz nicht in der Substanz, denn seine Pointe liegt gerade darin, daß Eigenschaften in letzter Analyse keine kausale Rolle spielen. Davidson unterscheidet auch in der Kausalitätstheorie strikt zwischen der rein extensionalen Kausalrelation und der von begrifflichen Mitteln abhängigen intensionalen Kausalerklärung. Die Kausalrelation besteht ausschließlich zwischen Ereignissen als reinen Individuen. Davidson macht den Punkt unmißverständlich klar:

„Given this extensionalist view of causal relations, it literally makes no sense, to speak of an event causing something as mental ... or as described in any other way.“²⁵

Kann man sagen, daß eine so verstandene Tokenidentitätstheorie das Problem der mentalen Verursachung wirklich löst? Nach meiner Einschätzung könnte man sagen, daß sie das Problem „auflöst“. Aber um welchen Preis? Die Intelligibilität der Identitätstheorie ist zumindest fragwürdig. Zum Beispiel macht gerade die ontologische Neutralität der rein partikulären Ereignisse die Kernthese unüberprüfbar, nach der sich die beiden Beschreibungen (mental und physisch) wirklich auf dieselben Ereignisse beziehen. Das eigentlich relevante metaphysische Faktum, auf dem die ganze Theorie aufbaut, liegt jenseits aller Beschreibungen und damit auch jenseits unseres epistemischen Zugriffs. Es handelt sich um ein metaphysisches Postulat. Davidsons notorische Schwierigkeiten, Identitätskriterien für Ereignisse anzugeben, sind nur ein Symptom dieser tieferliegenden Problematik. Außerdem: die Möglichkeit mentaler Verursachung bleibt – wie von vielen Autoren herausgearbeitet – unerklärt. Wenn wir fragen, warum hat Ereignis e_1 , das nachfolgende Ereignis e_2 verursacht, so fragen wir nach den kausal relevanten Eigenschaften von e_1 . Eine Identitätstheorie, die sich nur auf Ereignisse als rein partikuläre Individuen jenseits aller begrifflichen Beschreibung bezieht, hat auf diese zentrale Frage keine überzeugende Antwort.

Das Kernproblem liegt also in der rigorosen Trennung der metaphysi-

²⁵ D. Davidson, „Thinking Causes“, in: *Heil*, (Anm. 4) 3–17, hier 13.

schen von der epistemischen Ebene. In der Welt selbst treten unsere metaphysischen Probleme gar nicht auf, während wir mit unseren bescheidenen epistemischen Ressourcen in letzten unüberbrückbaren begrifflichen Dichotomien steckenbleiben. Nagel hat nicht zu Unrecht behauptet, daß man eine derartige Auffassung auch als Zwei-Aspekte-Theorie betrachten könne²⁶. Die Wirklichkeit präsentiert sich uns so, daß wir sie in zwei aufeinander irreduziblen begrifflichen Schemen beschreiben können. Dahinter liegt jedoch eine metaphysische Einheit, die sich nicht darum kümmert, wie wir sie beschreiben. Nagel selbst sieht die Grenzen dieser Lösung des Leib-Seele-Problems: „it has the faintly sickening odor of something put together in the metaphysical laboratory.“²⁷ Wenn meine Analysen korrekt sind, dann wurde bisher kein überzeugendes metaphysisches Modell entwickelt, das die Negation von Prinzip [2] plausibel machen könnte. Es folgt dann: Eine Lösung oder Auflösung des Problems der mentalen Verursachung im Sinne des metaphysischen Realismus stellt uns vor die Alternative, entweder dem Mentalen jede kausale Relevanz abzuspochen (Negation von Prinzip [3]) oder aber die Geschlossenheit des physischen Bereichs aufzugeben (Negation von Prinzip [1]). Eine dritte Möglichkeit scheint es nicht zu geben. Zum Abschluß möchte ich wenigstens andeuten, worin meines Erachtens eine solche dritte Möglichkeit bestehen könnte, die eine Negation von Prinzip [2] erlaubt, ohne in die genannten Schwierigkeiten zu geraten.

Verzicht auf metaphysischen Realismus als Ausweg?

Der Versuch, die unsere naturwissenschaftlichen Unternehmungen leitenden Überzeugungen mit den Überzeugungen, die uns als handelnde Wesen leiten, in einem einheitlichen Weltbild spannungsfrei zu harmonisieren, hat die gegenwärtige analytische Philosophie wenigstens zum Teil zu philosophischen Strategien zurückgeführt, die von den Gründervätern der analytischen Bewegung wegen ihrer metaphysischen Spekulationen als epistemologisch höchst problematisch bezeichnet worden wären. Vermutlich läßt sich der metaphysische Impuls im philosophischen Diskurs nicht ungestraft unterdrücken. Wenn aber metaphysische Fragen wiederaufgenommen werden, dann sollte der enge Zusammenhang zwischen Semantik und Ontologie, zwischen Epistemologie und Metaphysik, der vor allem von den Kritikern des sogenannten metaphysischen Realismus herausgearbeitet wurde, den Ausgangspunkt der Fragestellung markieren. Eine solche Grundlegung ontologischer Theorien auf einer semantischen Basis erzwingt keineswegs deflationäre philosophische Strategien. Es geht zunächst nur um den richtigen, methodisch reflektierten Ansatzpunkt. Dummett hat dies als die „bottom-up“ Vorgehensweise bezeichnet, die im Gegensatz zur klassischen

²⁶ *Th. Nagel*, *The View from Nowhere*, Oxford 1986, 30.

²⁷ *Ebd.* 49.

„top-down“ Methode nicht mit metaphysischen, sondern mit semantischen Untersuchungen beginnt.

Grundlegend für die Analyse des Leib-Seele-Problems wäre dementsprechend die wahrheitstheoretische und epistemologische Frage, ob und in welchem Sinne zwei Aussagen in zwei völlig verschiedenen und eventuell inkommensurablen Begriffssystemen (mental – physisch) beide wahr sein können. Die von Kritikern des metaphysischen Realismus immer wieder explizierte Möglichkeit mehrerer wahrer Versionen oder Welten ist für das philosophische Verständnis des Leib-Seele-Problems von zentraler Bedeutung. Es geht darüber hinaus vor allem auch um den semantischen Status von Kausalaussagen und Erklärungen. Man wird hier von der grundsätzlichen Relativität von Kausalaussagen auf epistemisch und pragmatisch relevante Kontexte ausgehen müssen. Die metaphysische Vorstellung einer die ganze Welt mit absoluter Notwendigkeit durchziehenden Kraft ist weit mehr, als zur semantischen Analyse von Kausalaussagen benötigt wird. Beide Einsichten – die Anerkennung einer Pluralität von Begriffswelten und die Beachtung der Kontext- oder Paradigmenabhängigkeit von Kausal-erklärungen – zusammen geben die Mittel an die Hand, um das Problem der mentalen Verursachung zu lösen.

Wenn der Begriff der „Welt an sich“ keinen Sinn ergibt, d. h. wenn Welten nur relativ auf sprachliche Ausdrucksmittel zu definieren sind, dann kann die im mentalistischen Vokabular beschriebene Wirklichkeit genau deshalb eine eigene Welt darstellen, weil sie irreduzibel ist auf die im physikalischen Vokabular beschriebene Wirklichkeit. Eventuell sind die beiden Welten sogar vollständig inkommensurabel. Es handelt sich um einen radikalen Wechsel des Themas, der den Charakter eines Bruches aufweist. McGinn hat in den letzten Jahren (in der Tradition Kants stehend) wiederholt argumentiert, daß dieser Bruch durch eine spezifische Begrenzung unseres Erkenntnisapparates bewirkt werde, der uns nur zwei vollständig verschiedene Erkenntniswege eröffne (mentale Introspektion des Bewußtseinsstromes und äußere Beobachtung)²⁸. Aus anti-realistischer Sicht ist es für Menschen unerreichbar (und daher sinnlos), hinter dieser Dualität der Erfahrung noch eine metaphysische Einheit jenseits all unserer epistemischen Kapazitäten erkennen zu wollen. Erkenntnis und Wahrheit nach Menschenmaß bleibt perspektivisch.

Eine ähnliche Argumentation läßt sich bezüglich des Kausalitätsbegriffs führen. In den Anfängen der analytischen Philosophie herrschte ein im weitesten Sinne Humesches Kausalitätsverständnis vor. Im Zuge der Entwicklung physikalistischer Ontologien mit Totalerklärungsanspruch entstand erneut das Bedürfnis nach einer metaphysischen Kausalitätstheorie, um so beispielsweise das Problem der Bezugnahme von Wörtern auf Gegenstände zu lösen. Allerdings wurden die mannigfaltigen epistemologischen Ein-

²⁸ Vgl. z. B. C. McGinn, „Can We Solve the Mind-Body Problem?“, in: *Mind* 98 (1989) 349–366.

wände gegen einen metaphysischen Kausalitätsbegriff keineswegs entkräftet. Putnam und andere haben sie in neuer Gestalt wieder vorgetragen²⁹. Der enge Zusammenhang zwischen Kausalaussagen mit der Theorie von Erklärungen wurde thematisiert. Ein kontextunabhängiger Kausalitätsbegriff ergibt kaum einen Sinn und ist explanatorisch nicht verwertbar, weil Kausalaussagen Erklärungen sind und Erklärungen (wie beispielsweise van Fraassen gezeigt hat) immer kontextrelativ sind³⁰.

Auf diesem Hintergrund ist die Lösung des Problems der Exklusivität von vollständigen Kausalerklärungen greifbar nahe. Paradigmen oder begriffliche Welten können ohne kausales Exklusionsproblem koexistieren, weil der Bezug auf die eine Welt (die einzige wahre Kausalverbindung) aus anti-realistischer Perspektive keinen Sinn ergibt. Jeder Begriffswelt kann dementsprechend auch eine eigene Weise der Kausalerklärung (ein Erklärungskontext) entsprechen. Zwischen der mentalen und der physikalischen Erklärung tritt kein Konkurrenzverhältnis mehr auf. Diese Lösung scheint vielleicht auf den ersten Blick trivial. Aber es sei hervorgehoben, daß eine Zwei-Erklärungen-Theorie nur unter Aufgabe des metaphysischen Realismus eine sinnvolle Option zur Lösung des Problems mentaler Verursachung darstellt. Ansonsten ist die Dichotomie zwischen epistemisch-pragmatischer Ebene (Pluralität der Erklärungen) und der metaphysischen Ebene (die einzig wahre Kausalordnung) nicht eliminierbar. Darum genügt auch nicht der Hinweis der Sprachdualisten auf eine faktische Pluralität der Begriffswelten oder Sprachspiele in der Umgangssprache. Eine erheblich größerer Aufwand an philosophischer Theorie ist gefordert. Erst wenn die Unterscheidung zwischen der epistemisch-pragmatischen und der metaphysischen Ebene mit guten philosophischen Gründen als unintelligibel klassifiziert wurde, ergibt sich die Möglichkeit, das Problem der mentalen Verursachung über eine Zwei-Erklärungen-Theorie zu lösen.

Die anti-realistische Option ist also nicht billig zu haben. Sie verlangt sogar einen noch höheren Preis, da sie unsere metaphysischen Fragen in einem spezifischen Sinne für unbeantwortbar hält. Unser metaphysischer Impuls hingegen drängt uns, das Problem der mentalen Verursachung so zu lösen, daß ein objektives metaphysisches Weltbild gewonnen werden kann. Wenn Philosophen diesem Impuls wirklich nachgeben sollen, wenn es tatsächlich die Aufgabe der Philosophie ist, metaphysische Probleme zu *lösen*, dann verbleibt als Fazit dieser Untersuchung, daß sich eine Lösung des Problems mentaler Verursachung in diesem Sinne in der gegenwärtigen Debatte noch immer nicht abzeichnet.

²⁹ Vgl. z. B. H. Putnam, „Is the Causal Structure of the Physical Itself Something Physical?“, in: *Ders., Realism with a Human Face*, Cambridge, MA 1990, 80–95.

³⁰ B. van Fraassen, *The Scientific Image*, Oxford 1980, bes. 97–157.

Brauchen wir ein neues Weltethos?

Universale Ethik in einer geschichtlichen Welt

VON EBERHARD SCHOCKENHOFF

Die Universalität und unbedingte Geltung ethischer Gebote wird gegenwärtig vor allem von zwei Seiten nachdrücklich in Erinnerung gerufen. Papst Johannes Paul II., der vor einigen Jahren die Religionsführer der Welt zum gemeinsamen Friedensgebet nach Assisi einlud, erneuert in seiner jüngsten Enzyklika „Veritatis Splendor“ die Ablehnung jedes ethischen Relativismus durch das kirchliche Lehramt und betont: „Insofern aber das Naturgesetz die Würde der menschlichen Person zum Ausdruck bringt und die Grundlage für ihre fundamentalen Rechte und Pflichten legt, ist es in seinen Geboten universal, und seine Autorität erstreckt sich über alle Menschen.“¹ Fast zeitgleich zum Erscheinen der päpstlichen Moralencyklika – im Spätsommer des Jahres 1993 – tagt in Erinnerung an seine erste Zusammenkunft vor hundert Jahren in Chicago das „Parlament der Weltreligionen“, um die Teilnehmer aus über 250 Glaubensgemeinschaften auf die Annahme eines gemeinsamen Weltethos zu verpflichten. Dies geschieht ebenfalls mit der ausdrücklichen Zielsetzung, die „Unbedingtheit und Universalität ethischer Verpflichtung“, die in der sittlichen Autonomie des Menschen wohl angelegt, aber noch nicht letztgültig verankert sei, zu begründen².

Während der Papst den universalen Anspruch der sittlichen Wahrheit durch den Gedanken erläutert, die Würde des Menschen schließe um der Möglichkeit ihrer authentischen Verwirklichung willen „wesentlich auch die Achtung einiger Grundgüter ein, ohne deren Respektierung man dem Relativismus und der Willkür verfällt“³, verfolgt das Projekt Weltethos nach Methode und inhaltlicher Stoßrichtung einen anderen Weg. Es versucht, durch einen Vergleich der vorhandenen Ethosbestände aller Religionssysteme zwar keine numerische Einstimmigkeit, wohl aber moralische Einmütigkeit über einen „minimalen Grundkonsens bezüglich bestimmter Werte, Normen und Haltungen“ zu erzielen, ohne den ein friedliches Zusammenleben in der Weltgemeinschaft der Völker unmöglich ist⁴. Die Konvergenz grundlegender ethischer Überzeugungen soll so aufgedeckt, bewußtgemacht und durch interreligiösen Dialog und weltweite ökumenische

¹ Enzyklika „Veritatis Splendor“, Nr. 51.

² H. Küng, Auf der Suche nach einem universalen Grundethos der Religionen, in: Conc(D) (1990) 154–164, hier: 160.

³ Wie Anm. 1: Nr. 48.

⁴ Auf der Suche, 155 f., vgl. auch *ders.*, Geschichte, Sinn und Methode der Erklärung zu einem Weltethos, in: *Ders.*, K.-J. Kuschel (Hg.), Erklärung zum Weltethos. Die Deklaration des Parlamentes der Weltreligionen, München 1993, 49–87, 67. Eine sorgfältige, dokumentierte Darstellung findet sich bei F. Noichl, Das „Projekt Weltethos“ aus moraltheologischer Sicht (erscheint im Jahrbuch für Religionswissenschaft 1994).