



Prof. Dr. Christian Nimtz
www.nimtz.net // lehre@nimtz.net

Philosophie des Geistes

Kapitel 6: Einwände gegen die Identitätstheorie II: Kripkes Argument und das Problem der Multirealisierbarkeit

-1-



Literatur

- ★ Saul Kripke: Identity and Necessity, in: Milton Munitz (ed.) (1971): Identity and Individuation, New York: New York University Press, 135–164.
- ★ Jerry Fodor: Special sciences, or the Disunity of Science as a Working Hypothesis, in: Synthèse 28, 77–115, 1974.

Ansgar Beckermann 2008: Analytische Einführung in die Philosophie des Geistes, Berlin: de Gruyter, Kapitel 5.

Jerry Fodor: Special Sciences. Still Autonomous after all these Years, in: Nous 31, 1997, 149–163.

Saul Kripke: Identity and Necessity, in: Naming and Necessity, Oxford 1980, lecture 3, bes. 144-155.

-2-



Das Programm des Kapitels

- §1 Notwendigkeit und A Priorizität
- §2 Kripkes Argument gegen die Identitätstheorie
- §3 Das Problem der Multirealisierbarkeit mentaler Zustände
- §4 Fodors antireduktionistisches Wissenschaftsbild
- §5 Konsequenzen für die Identitätstheorie

-3-



§1 Notwendigkeit und A Priorizität

Einwand: Die Beziehung zwischen mentalen und Hirnzuständen ist doch bestenfalls kontingent.

Smarts Antwort: Das zeigt nur, dass mentale und physikalische Prädikate sinnverschieden sind, nicht aber, dass diese Prädikate nicht dieselbe Eigenschaft bezeichnen.

Voraussetzung: Die Beziehung zwischen zwei Eigenschaften *M* und *P* ist genau dann **notwendig** (nicht kontingent), wenn der Satz „Für alle *x*: *x* hat *M* genau dann, wenn *x* *P* hat“ *analytisch wahr* ist.

-4-



- ➔ Smart geht davon aus, dass die drei Begriffe **notwendig**, **analytisch** und **a priori wahr** extensionsgleich sind, und es gilt:
 x ist notwendig wahr \leftrightarrow x ist analytisch wahr \leftrightarrow x ist a priori wahr

Dagegen argumentierte Saul Kripke in seiner Vorlesung „Naming and Necessity“ von 1970.

Immerhin ist Notwendigkeit eine **metaphysische** und A Priorizität eine **epistemische** Kategorie – und die können laut Kripke sehr wohl auseinander fallen.



Kripke zufolge müssen wir vier Arten von Wahrheiten unterscheiden:

- notwendige Wahrheiten a priori
- kontingente Wahrheiten a posteriori
- notwendige Wahrheiten a posteriori
- kontingente Wahrheiten a priori

Dies ist für die Identitätstheorie von größter Bedeutung, weil nach Kripkes Überlegungen die *a posteriori* wahren Aussagen nicht mehr einfach mit den *kontingenterweise* wahren Aussagen gleichgesetzt werden können.



Kripke ist der Überzeugung, dass a posteriori wahre Identitätsaussagen wie „Mount Everest = Tschomolungma“ oder „Die Eigenschaft, ein Blitz zu sein = die Eigenschaft, eine bestimmte Art elektrischer Entladung zu sein“ **notwendig** wahr sind.

Dabei gilt: Eine Aussage A ist genau dann **notwendig** wahr, wenn A in allen möglichen Welten wahr ist, d.h. wenn es keine denkbare Situation gibt, in der A falsch ist. (Eine mögliche Welt ist eine vollständige Art und Weise, wie sich die Dinge verhalten könnten.)

- ➔ Wenn Kripke recht hat, würde der Nachweis, dass es für jeden mentalen Zustand M und jeden physischen Zustand P zumindest (metaphysisch) möglich ist, dass jemand in M, aber nicht in P ist (oder umgekehrt), die Identitätstheorie zu Fall bringen.



(K1) Alle wahren Identitätsaussagen der Form „a = b“ sind notwendig wahr, wenn die Ausdrücke a und b starre Bezeichner sind.

Ein **starrer Bezeichner** (*rigid designators*) ist ein Bezeichner, der in jeder möglichen Welt denselben Gegenstand bezeichnet.

- ◆◆ So gesehen ist die These (K1) trivial; wenn die Aussage „a = b“ wahr ist und ‚a‘ und ‚b‘ in jeder möglichen Welt denselben Gegenstand bezeichnen, dann muss „a = b“ offenbar in jeder möglichen Welt wahr sein.

Kripke argumentierte dafür, dass Eigennamen und Eigenschaftsbezeichner starre Bezeichner sind.

§1 Notwendigkeit und A Priorizität



Wahre Identitätsaussagen wie diese sind nicht notwendig wahr:

- Benjamin Franklin = der Erfinder der Zwei-Stärken-Brille.
- Die Venus = der hellste Stern am Abendhimmel.

Denn Ausdrücke wie ‚der Erfinder der Zwei-Stärken-Brille‘ sind kein starrer Bezeichner. Die Zwei-Stärken-Brille hätte nämlich auch von Spinoza erfunden werden können. Also gibt es eine mögliche Welt, in welcher der Ausdruck ‚der Erfinder der Zwei-Stärken-Brille‘ nicht auf Benjamin Franklin, sondern auf Spinoza zutrifft.

Analoges gilt für ‚der hellste Stern am Abendhimmel‘

-9-

§1 Notwendigkeit und A Priorizität



Wie steht es mit den Begriffen ‘a priori’ und ‘a posteriori’?

Traditionell wird ein Satz ‘a priori wahr’ genannt, wenn seine Wahrheit ohne Rückgriff auf Erfahrung eingesehen werden kann; ‘a posteriori wahr’ dagegen heißt ein Satz, wenn er nur unter Bezug auf Erfahrung gerechtfertigt werden kann.

Diese Erläuterung leidet allerdings daran, dass nicht klar ist, was hier unter ‘Erfahrung’ verstanden werden soll. Nur Sinneserfahrung oder auch Introspektion oder vielleicht sogar das, was manchmal ‘Intuition’ genannt wird?

-10-

§1 Notwendigkeit und A Priorizität



Trotzdem ist der folgende Satz offenbar *a priori* wahr:

- (1) Alle Junggesellen sind unverheiratet

Denn um seine Wahrheit einzusehen, benötigt man nur das linguistische Wissen, dass der Ausdruck ‘Junggeselle’ auf (und nur auf) unverheiratete männliche Erwachsene zutrifft. Wer der deutschen Sprache mächtig ist, weiß also schon allein aufgrund seiner Sprachkompetenz, dass der Satz (1) wahr ist.

-11-

§1 Notwendigkeit und A Priorizität



Dasselbe scheint für folgenden Satz zu (2) „Ich bin jetzt hier“ zu gelten. Die Regeln der deutschen Sprache besagen doch:

- Der Ausdruck ‘ich’ bezeichnet den jeweiligen Sprecher.
- Der Ausdruck ‘jetzt’ bezeichnet den jeweiligen Äußerungszeitpunkt.
- Der Ausdruck ‘hier’ bezeichnet den jeweiligen Äußerungsort.

Jeder, der den Satz (2) äußert, scheint daher etwas Wahres zu sagen; denn jeder ist in dem Moment, in dem er einen Satz äußert, an dem Ort, an dem er den Satz äußert.

- **Generell:** Ein Satz ist genau dann *a priori* wahr, wenn man ohne Rückgriff auf Erfahrung einsehen kann, dass jeder etwas Wahres sagt, der diesen Satz äußert.

-12-

§1 Notwendigkeit und A Priorizität



Ein Satz ist

notwendig wahr genau dann, wenn er in allen möglichen Welten wahr ist.

kontingent wahr genau dann, wenn er wahr, aber nicht in allen möglichen Welten wahr ist.

a priori wahr genau dann, wenn man ohne Rückgriff auf Erfahrung einsehen kann, dass jeder etwas Wahres sagt, der diesen Satz äußert.

a posteriori wahr genau dann, wenn er wahr, aber nicht *a priori* wahr ist.

-13-

§1 Beispiele für notwendige Wahrheiten a posteriori



- (3) Cicero = Tullius
- (4) Der Morgenstern = der Abendstern,
- (5) Wasser = H₂O.

Diese Aussagen sind notwendig wahr, weil alle Bezeichner ihn ihnen starr sind. Aber sie sind nur *a posteriori* wahr.

Aufgrund der Sinnverschiedenheit der Bezeichner reicht linguistisches Wissen allein nicht aus, um zu erkennen, dass jeder etwas Wahres sagt, der einen dieser Sätze äußert.

Nur aufgrund von Erfahrung können wir wissen, dass die betreffenden Bezeichner jeweils dieselbe Person, denselben Gegenstand bzw. denselben Stoff bezeichnen.

-14-

§1 Ein Beispiel für eine kontingente Wahrheit a priori



- (2) Ich bin jetzt hier.

Wir hatten schon gesehen, dass dieser Satz *a priori* wahr ist. Denn allein aufgrund unserer Sprachkompetenz wissen wir, dass jeder, der diesen Satz äußert, etwas Wahres sagt.

Aber keine seiner Äußerungen ist notwendig wahr. Denn natürlich könnte jeder, der diesen Satz äußert, zu dem Zeitpunkt, an dem er ihn äußert, woanders sein.

D.h., natürlich gibt es viele mögliche Welten, in denen sich derjenige, der diesen Satz äußert, zum Zeitpunkt seiner Äußerung an einem anderen Ort aufhält als dem, an dem er sich zu diesem Zeitpunkt in der wirklichen Welt befindet.

-15-

§2 Kripkes Argument gegen die Identitätstheorie



Was folgt aus all dem für die Identitätstheorie? Zunächst einmal, dass eine Identitätsaussage wie

- (6) Schmerz ist identisch mit dem Feuern von C-Fasern

nur dann wahr sein kann, wenn sie notwendig wahr ist – vorausgesetzt, dass die Ausdrücke 'Schmerz' und 'das Feuern von C-Fasern' starre Bezeichner sind.

- ♦♦ Kripke will zeigen, dass die Aussage (6) nicht notwendig wahr ist und dass wir daher – da die Ausdrücke 'Schmerz' und 'das Feuern von C-Fasern' starre Bezeichner sind – um die Schlussfolgerung nicht herumkommen, dass (6) sogar falsch ist.

-16-



Erstens: ‚Schmerz‘ und ‚das Feuern von C-Fasern‘ sind starre Bezeichner

„Tatsächlich scheint es, dass beide Ausdrücke ‚mein Schmerz‘ und ‚mein sich-in-dem-und-dem-Gehirnzustand-Befinden‘ ... starre Bezeichner sind. Das heißt, wann immer etwas der und der Schmerz ist, dann ist es wesentlich genau dies, nämlich der und der Schmerz, und wann immer etwas der und der Gehirnzustand ist, dann ist es wesentlich genau dies, nämlich der und der Gehirnzustand. Somit sind beide Ausdrücke starre Bezeichner. Man kann nicht sagen, dieser Schmerz hätte etwas anderes sein können, ein anderer Zustand. Beide Ausdrücke sind starre Bezeichner.“ (Kripke 1971, 213f.)



Zweitens: Die Aussage (6) ist nicht notwendig wahr

Das zentrale Argument für seine These, dass die Aussage (6) nicht notwendig wahr ist, ergibt sich für Kripke aus der Tatsache, dass sich herausstellen könnte, dass Schmerz nicht mit dem Feuern von C-Fasern identisch ist; d.h. daraus, dass der Satz wahr ist:

- (6') Es könnte sich herausstellen, dass Schmerz nicht mit dem Feuern von C-Fasern identisch ist.

Denn seiner Meinung nach gilt:

- (7) Wenn sich herausstellen kann, dass p , dann ist es auch möglich, dass p .



Die Form des Arguments

- (i) Schmerz = das Feuern von C-Fasern
- (ii) Wenn die Aussage „Schmerz = das Feuern von C-Fasern“ wahr ist, dann muss sie auch notwendig wahr sein.
- (iii) Es ist möglich, dass gilt: Schmerz \neq das Feuern von C-Fasern.
- (iv) Also: Schmerz \neq das Feuern von C-Fasern.



Dieses Argument scheint jedoch eine Schwäche zu haben. Denn Kripke selbst schreibt häufig, *dass sich herausstellen könnte*, dass Wasser nicht H₂O ist, dass Wärme nicht die Bewegung von Molekülen ist oder dass Blitze keine elektrischen Entladungen sind. Auf der anderen Seite ist er jedoch zugleich von der notwendigen Wahrheit dieser Aussagen überzeugt. Also scheint Kripke selbst an der Richtigkeit des Prinzips (7) zu zweifeln.

§2 Kripkes Argument gegen die Identitätstheorie



Kripkes Erwiderung: Der Satz

(8') Es könnte sich herausstellen, dass Wasser nicht H_2O ist
ist genau genommen falsch. Wahr ist nur die verwandte Aussage:

(8'') Es ist möglich, dass sich überall dort, wo in unserer Welt tatsächlich Wasser ist (in Seen und Flüssen, im Meer, im Schwimmbecken, usw.) ein *anderer* Stoff befindet, der dieselben Oberflächeneigenschaften hat wie Wasser – d.h. der genauso aussieht, schmeckt und riecht wie Wasser, der genauso unseren Durst löscht wie Wasser, usw.

-21-

§2 Kripkes Argument gegen die Identitätstheorie



Aus der Wahrheit *dieser* Aussage folgt aber nicht, dass diese Aussage nur kontingent und eben deshalb falsch ist:

(8) Wasser ist H_2O

Mit der Aussage (6) ist das seiner Meinung nach aber anders. Denn wenn wir versuchen, zu der Aussage

(6') Es könnte sich herausstellen, dass Schmerz nicht das Feuern von C-Fasern ist

eine Aussage zu formulieren, die der Aussage (8'') entspricht, kommen wir zu:

(6'') Es ist möglich, dass wir immer dann, wenn wir sagen, dass wir Schmerzen haben, in einem Zustand sind, der gar kein Schmerz ist, der sich aber genauso schmerzhaft anfühlt wie richtige Schmerzen.

-22-

§2 Kripkes Argument gegen die Identitätstheorie



Dies ist aber Unsinn!

Denn es ist zwar möglich, dass ein anderer Stoff als Wasser dieselben Oberflächeneigenschaften hat wie Wasser (dass er genauso aussieht, schmeckt und riecht wie Wasser, dass er genauso unseren Durst löscht wie Wasser, etc.).

Aber es ist nicht möglich, dass ein Zustand, der kein Schmerz ist, sich genauso anfühlt wie Schmerz. Denn Schmerzen **sind** die Zustände, die sich schmerzhaft anfühlen. Jeder Zustand, der sich genauso anfühlt wie Schmerz, ist ein Schmerz.

➔ Die Aussage (6) ist also nicht bloß scheinbar, sie ist wirklich kontingent. Und daher muss sie falsch sein. Denn wahr kann sie nur sein, wenn sie notwendig wahr ist.

-23-

§3 Das Problem der Multirealisierbarkeit mentaler Zustände



Eine Identitätsaussage wie (1) kann nur dann wahr sein, wenn der Satz (2) zumindest ein *wahres Naturgesetz* ist:

- (1) Schmerz = das Feuern von C-Fasern
- (2) Ein Wesen hat genau dann Schmerzen, wenn seine C-Fasern feuern

Problem: Wir wissen heute, dass ein bestimmter mentaler Zustand bei *verschiedenen Personen* durchaus mit unterschiedlichen neuronalen Zuständen korreliert sein kann.

-24-

§3 Das Problem der Multirealisierbarkeit mentaler Zustände



Wenn man Personen eine bestimmte Aufgabe vorlegt und zugleich z.B. mittels eines PET-Scans die Aktivierungsmuster ihrer Gehirne beobachtet, ergeben sich in der Regel zwar sehr ähnliche, aber nicht exakt dieselben Muster.

Wir wissen weiter, dass sich sogar bei ein und derselben Person die Korrelation zwischen mentalen und Gehirnzuständen im Laufe ihres Lebens dramatisch verändern kann.

Nach Gehirnverletzungen z.B. können andere Teile des Gehirns die Funktionen des geschädigten Gewebes übernehmen. Nur aufgrund dieser Plastizität des Gehirns sind wir in der Lage, im Laufe unseres Lebens unsere geistigen Fähigkeiten trotz des täglichen Zugrundegehens tausender von Nervenzellen (einigermaßen) zu erhalten.

-25-

§3 Das Problem der Multirealisierbarkeit mentaler Zustände



Wir wissen schließlich, dass sich die Neurophysiologie der meisten Tiere mehr oder weniger stark von der unsrigen unterscheidet.

Soll allein daraus schon folgen, dass diese Tiere nicht dieselben mentalen Zustände haben wie wir?

Und wie steht es schließlich mit Marsmenschen und Robotern? Sollen diese Wesen schon deshalb kein dem unseren vergleichbares mentales Leben haben, weil ihr 'Gehirn' nicht aus Nervenzellen, sondern z.B. aus Silizium-Chips besteht?

- ♦♦ Offensichtlich machen es alle diese Überlegungen ziemlich unwahrscheinlich, dass jeder mentale Zustand mit genau einem neuronalen Zustand korreliert ist. Und eben deshalb steht die Identitätstheorie empirisch auf einer äußerst schwachen Grundlage.

-26-

§4 Fodors antireduktionistisches Wissenschaftsbild



Jerry Fodor hat in seinem Aufsatz „Special Sciences“ (1974) dieses Ergebnis generalisiert

Fodors These

Es ist im allgemeinen **nicht** der Fall, dass den Artbegriffen, die in Einzelwissenschaften wie der Geologie, Psychologie, Soziologie oder der Wirtschaftswissenschaft eine entscheidende Rolle spielen, in eindeutiger Weise Artbegriffe der Physik entsprechen.

Man kann daher grundsätzlich nicht davon ausgehen, dass diese Wissenschaften auf die Physik reduzierbar sind.

-27-

§4 Fodors antireduktionistisches Wissenschaftsbild



„Die Gründe dafür, dass es unwahrscheinlich ist, dass jede Art einer physikalischen Art entspricht, sind: (a) Häufig kann man interessante Generalisierungen (z.B. Generalisierungen, die kontrafaktische Aussagen stützen) über Ereignisse formulieren, deren physikalische Beschreibungen keine Gemeinsamkeiten aufweisen. (b) Oft ist die Frage, *ob* die physikalischen Beschreibungen der unter solche Generalisierungen subsumierten Ereignisse etwas gemeinsam haben, offensichtlich vollkommen irrelevant für die Wahrheit dieser Generalisierungen oder für ihre Interessanztheit oder für den Grad ihrer Bestätigung oder überhaupt für irgendeine ihrer epistemologisch wichtigen Eigenschaften.“ ...

-28-

§4 Fodors antireduktionistisches Wissenschaftsbild



(c) Die Einzelwissenschaften beschäftigen sich vorwiegend mit der Formulierung von Generalisierungen dieser Art. Ich gehe davon aus, dass diese Bemerkungen offensichtlich sind bis hin zur Selbstbestätigung; sie springen ins Auge, sobald man den (anscheinend radikalen) Schritt macht, die Existenz der Einzelwissenschaften überhaupt ernst zu nehmen.“ (Fodor 1974, 141 – die Übersetzung weicht z.T. von der deutschen Ausgabe ab)

-29-

§4 Ein Beispiel



Betrachten wir den wirtschaftswissenschaftlichen Begriff *Austausch von Zahlungsmitteln*, dann ist es Fodor zufolge durchaus wahrscheinlich, dass jedes einzelne Ereignis, das unter diesen Begriff fällt, d.h. dass jeder einzelne Austausch von Zahlungsmitteln ein physikalisches Ereignis in dem Sinne ist, dass es für dieses Ereignis eine wahre Beschreibung im Vokabular der Physik gibt.

Unwahrscheinlich ist jedoch, dass alle diese Ereignisse unter denselben physikalischen Artbegriff fallen.

-30-

§4 Ein Beispiel



Ein Austausch von Zahlungsmitteln z.B. mag darin bestehen,

- dass auf Schnüre aufgereichte Schnecken- und Muschelscheiben (Wampums) den Besitzer wechseln;
- dass 10 Euro-Scheinen den Besitzer wechseln
- dass ein Scheck unterschreiben und überreicht wird; usw.

Es wäre daher schon recht merkwürdig, wenn allen diesen verschiedenen Tätigkeiten etwas **physikalisch** Interessantes gemeinsam wäre.

-31-

§4 Der entscheidende Punkt



Die verschiedenen Arten des Zahlungsverkehrs haben interessante Gemeinsamkeiten und werden deshalb von den Wirtschaftswissenschaften unter denselben Begriff „Austausch von Zahlungsmitteln“ subsumiert.

Physikalisch gesehen haben die verschiedenen Arten des Zahlungsverkehrs aber keine Gemeinsamkeiten. (Und wenn sie sie hätten, wäre das irrelevant.)

- Das Bild, das reduktionistisch eingestellte Wissenschaftler vom Verhältnis der Einzelwissenschaften zur Physik haben, muss grundlegend revidiert werden.

-32-

§4 Das traditionelle Bild



Wenn wir in einer Einzelwissenschaft E ein Gesetz haben der Form

- (i) Für alle x : Wenn x F hat, hat x auch P ,

dann entsprechen den Begriffen F und P von E physikalische Begriffe F und P , so dass sich aus der Physik das Gesetz

- (ii) Für alle x : Wenn x P hat, hat x auch F

ableiten lässt.

-33-

§4 Ein Beispiel



Dem Begriff der Temperatur T der klassischen Thermodynamik entspricht der physikalische Begriff der mittleren kinetischen Energie $mv^2/3k$, so dass zum Gesetz von Boyle und Charles (BC) aus der statistischen Mechanik das Bildgesetz (BC*) abgeleitet werden kann.

$$(BC) \quad P \cdot V = N \cdot k \cdot T$$

$$(BC^*) \quad P \cdot V = N \cdot k \cdot mv^2/3k$$

-34-

§4 Fodors antireduktionistisches Wissenschaftsbild



Die verschiedenen Ereignisse, die unter einen einzelwissenschaftlichen Artbegriff F fallen, können auf sehr unterschiedliche Weise physikalisch realisiert sein. D.h., dem Artbegriff F werden viele verschiedene physikalische Begriffe P_1, \dots, P_n, \dots entsprechen. Und das Verhältnis zwischen einzelwissenschaftlichen und physikalischen Gesetzen wird sich eher so darstellen

Angenommen, das Ereignis Fa verursacht das Ereignis Pa und Fa wird durch das physikalische Ereignis P_1a realisiert.

-35-

§4 Fodors antireduktionistisches Wissenschaftsbild

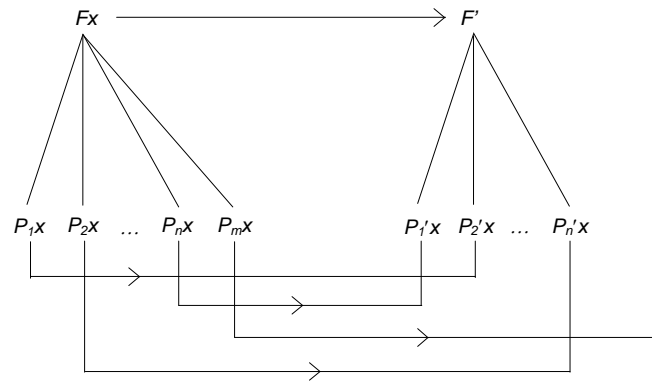


Dann gilt folgendes: P_1a verursacht ein anderes physikalisches Ereignis P_2a , das seinerseits eine Realisierung von Pa ist.

Allerdings sind da auch Ausnahmen möglich; d.h., es kann Realisierungen von Fa geben, die keine Realisierungen von Pa verursachen. Und dies ist der systematische Grund dafür, dass einzelwissenschaftliche Gesetze im allgemeinen nicht ausnahmslos gelten.

-36-

§4 Fodors antireduktionistisches Wissenschaftsbild



-37-

§5 Konsequenzen für die Identitätstheorie



- ◆◆ Es ist klar, dass die Identitätstheorie auf völlig hoffnungslosem Posten steht, wenn Fodor mit diesen Überlegungen recht hat. Und alles spricht dafür, dass er damit recht hat.
- ◆◆ Die Multirealisierbarkeit mentaler Zustände und Eigenschaften gilt daher heute als das entscheidende Argument gegen die Identitätstheorie.

-38-

- Ende -



-39-