

1. März 2020

UNIVERSITÄT BERN

---

**Yes, I did it!**

*Eine Replik auf Kims These der physikalischen  
Reduktionierbarkeit von mentalen Kausalerklärungen*

---

Arbeit zum Seminar *Das Problem der mentalen Verursachung* (455296-HS2019) von Prof. Dr. Vera Hoffmann-Kolss

Thomas König  
Matrikel #: 85-917-151  
thomas.koenig@students.unibe.ch

Dieser Text hat 5406 Wörter.

Ich danke Zineb Benkhelifa für das sorgfältige Gegenlesen.

## Inhalt

1	Einleitung	2
2	Rekonstruktion des Arguments von Kim	3
3	Relationale Eigenschaften und Kausalität des Mentalen	4
3.1	Mentale Zustände sind an sich relational . . . . .	4
3.1.1	Relation zum Subjekt . . . . .	4
3.1.2	Privater Zugang in Form von Qualia . . . . .	4
3.2	Leiblichkeit als mentale Relationen zum eigenen Körper . . . . .	5
3.3	Relationale Eigenschaften von Handlungen . . . . .	6
4	Ungleiche relationale Eigenschaften und Identität	7
5	Folgen für das Argument von Kim	7
5.1	Reduktionismus von relationalen Eigenschaften . . . . .	8
5.1.1	Zur Überbestimmtheit relationaler Eigenschaften . . . . .	9
5.1.2	Die Erklärungslücke . . . . .	10
6	Normative Relevanz von angemessenen kausalen Selbstattributionen	11
7	Diskussion	12
	Literaturverzeichnis	12

## 1 Einleitung

Wenn wir nach den Gründen für unser Tun gefragt werden, beginnen wir üblicherweise von Wünschen, Erwartungen, Bedürfnissen, Motiven, Intentionen, Überzeugungen und Ähnlichem zu sprechen. Wir begründen unser Tun also mit Begriffen aus der Domäne des Mentalen. Wenn aber das, was wir mit solchen Begründungen meinen, in der physischen Welt tatsächlich ursächlich wirksam sein soll, dann muss es entweder selbst gemäss den Regeln der Physik wirken, oder die Regeln der Physik sind nicht hinreichend für die Erklärung dessen, was in der Welt ursächlich möglich ist.

Beide möglichen Folgen des obigen Konditionals scheinen problematisch: Die erste Alternative legt nahe, dass wir unsere Erklärungen unseres Tuns in mentalen Begriffen im Prinzip durch Erklärungen allein anhand der Regeln der Physik vollständig ersetzen könnten. Mentale Erklärungen von Ursachen wären dann auf physische Erklärungen von Ursachen reduzierbar und daher überflüssig. Die zweite Alternative ist problematisch, weil sie mit dem gängigen, sehr erfolgreichen naturwissenschaftlichen Weltbild nicht zu vereinen scheint. Wie denn, wenn nicht als physische Ursache, soll den etwas in der physischen Welt überhaupt wirksam sein können? Zuletzt können wir natürlich das Antezedens des obigen Konditionals verneinen, was bedeuten würde, dass mentale Vorgänge gar nichts bewirken, also kausal inert und ein blosses Epiphänomen sind. Intuitiv ist keine dieser drei Alternativen wirklich attraktiv.

Die Position, die keine dieser drei Alternativen akzeptiert, wird als nicht-reduktiver Physikalismus von mentalen Ursachen bezeichnet. Nicht-reduktiver Physikalismus von mentalen Ursachen besagt, dass mentale Zustände kausal wirksam sind, weil sie durch physische Zustände realisiert werden, die kausal wirksam sind. Gleichzeitig postuliert die Position, dass mentale Zustände nicht auf physische Zustände reduziert werden können, also von physischen Zuständen verschieden sind. Die Position des nicht-reduktiven Physikalismus ist systematisch kritisiert worden. Insbesondere gibt es ein klassisches Argument von [Kim \[2005\]](#), das besagt, dass wenn die physische Realisierung eines mentalen Zustands die kausale Wirkung dieses Zustand erklären soll, die Erklärung so vollständig sein muss, dass nichts mehr bleibt, das durch den mentalen Zustand selbst zu erklären wäre. Daraus folgt für Kim, dass mentale Zustände keine kausale Wirkung haben können.

Es gibt bereits verschiedene Ansätze, die unterschiedliche Prämissen dieses Arguments angreifen [[Baker, 1993](#), [Bennett, 2003](#), [Horgan, 1989](#), [Kroedel and Schulz, 2016](#), [Woodward, 2008](#)]. Ich werde im Folgenden einen weiteren Ansatz verfolgen, aus dem sich ergeben soll, dass Kims Argument nicht schlüssig ist. Die These, die ich verfolgen werde ist, dass es nicht notwendigerweise problematisch ist, von einer Identität des Mentalen mit dem Physischen auszugehen und gleichzeitig an einem nichtreduktiven Physikalismus festzuhalten. Die Nicht-Identität des Mentalen mit dem Physischen ist, in Kims Formulierung, notwendiger Teil des nichtreduktiven Physikalismus und folgt in seiner Darstellung aus der Nichtreduzierbarkeit des Mentalen. Denn nur wenn sich das Mentale vom Physischen unterscheidet, lässt sich aufrechterhalten, dass sich das Mentale nicht auf das Physische reduzieren lässt. Die Nicht-Identität des Mentalen mit dem Physischen (als Folge der für den nicht-reduktiven Physikalismus notwendigen Unterscheidung des Mentalen vom Physischen), folgt in Kims Argument implizit aus den Leibnizschen Gesetzen der Identität des Nicht-Unterscheidbaren und der Nicht-Unterscheidbarkeit des Identischen. Was Kim hier aber ausser Acht lässt ist, dass die Leibnizschen Gesetze nicht gültig sind, wenn es um relationale Eigenschaften geht. Mein Argument wird darauf hinauslaufen, dass sich das Mentale vom Physischen notwendigerweise unterscheidet, obwohl es mit ihm a posteriori identisch ist, weil das Mentale notwendigerweise relationale Eigenschaften hat, die dem Physischen notwendigerweise fehlen. Diese relationalen Eigenschaften beziehen sich auf das Subjekt der mentalen Zustände, ohne das es mentale Zustände gar nicht gäbe.

Der Essay wird in einem ersten Teil das Argument von Kim darstellen. In einem zweiten Teil werde ich herausarbeiten, dass Ursachen, die in mentalen Begriffen verstanden werden

notwendigerweise relationale Eigenschaften haben. In einem dritten Teil werde ich auf bekannte Standardargumente zurückkommen, die zeigen, dass unterschiedliche relationale Eigenschaften nicht im Widerspruch stehen zur möglichen Identität dessen, was sich relational unterscheidet. Im vierten Teil werde ich zeigen, dass Kims Argument nicht mehr schlüssig ist, wenn wir davon ausgehen, dass mentale Ursachen mit physischen Ursachen identisch sind, sich aber in ihren relationalen Eigenschaften unterscheiden. Im fünften Teil werde ich schliesslich noch kurz darauf hinweisen, dass die für das Argument relevanten relationalen Eigenschaften gewissen Normen unterliegen, und dass Kims Rekonstruktion von mentalen Ursachen diese Normen nicht abbilden kann.

## 2 Rekonstruktion des Arguments von Kim

Nach Kim [Kim, 2005] ist die kausale Wirksamkeit von mentalen Zuständen inkompatibel mit der vollständigen Annahme der folgenden 4 Aussagen: (i) Mentale Zustände werden durch physische Zustände realisiert (Supervenienzprinzip, [Kim, 2005, S. 33]), (ii) die physische Welt ist in sich kausal abgeschlossen [Kim, 2005, S. 43], (iii) mentale Zustände sind nicht auf physische Zustände reduzierbar [Kim, 2005, S. 34] und (iv) unterschiedliche, aber jeweils hinreichende kausale Erklärungen für dasselbe Ereignis schliessen sich normalerweise aus (Exklusion) [Kim, 2005, S. 42].

- Die These der Verursachung von Ereignissen ( $E$ ) durch mentale Zustände ( $M$ ) wird als Prämisse vorausgesetzt.

$$M \rightarrow E \tag{1}$$

- Aussage (i): Der mentale Zustand  $M$  wird durch einen physischen Zustand  $P$  realisiert (ich verwende dafür das Symbol  $\mapsto$ ):

$$\exists P(P \mapsto M) \tag{2}$$

- Aussage (ii): Alle physischen Ereignisse müssen eine physische Ursache haben (Geschlossenheit der physikalischen Welt):

$$\forall E(\exists P(P \rightarrow E)) \tag{3}$$

- Aussage (iii): Aus der Nichtreduzierbarkeit von  $M$  auf  $P$  folgt gemäss Kim [2005, S. 42] die Identität des Mentalen mit dem Physischen:

$$M \neq P \tag{4}$$

- Aussage (iv): Ereignisse sind (üblicherweise) nicht kausal überbestimmt. Daraus folgt:

$$(M \neq P) \rightarrow \neg((M \rightarrow E) \wedge (P \rightarrow E)) \tag{5}$$

- Aus 4 und 5 folgt, dass das Ereignis  $E$  nicht gleichzeitig durch physische und mentale Ursachen erklärt werden kann (abhängig von den Prämissen 4 und 5):

$$\neg((M \rightarrow E) \wedge (P \rightarrow E)) \tag{6}$$

- Aus 1, 2 und 3 folgt: Die physische Realisierung  $P$  des mentalen Zustands  $M$  verursacht das Ereignis  $E$  (abhängig von den Prämissen 1, 2 und 3).

$$\exists P((P \mapsto M) \wedge (P \rightarrow E)) \tag{7}$$

- Aus 1, 2 und 7 folgt, dass das Ereignis  $E$  sowohl durch physische als auch durch mentale Ursachen erklärt wird (abhängig von den Prämissen 1, 2 und 3):

$$(M \rightarrow E) \wedge (P \rightarrow E) \quad (8)$$

- Aus 6 und 8 folgt ein Widerspruch (abhängig von den Prämissen 1, 2, 3, 4 und 5)

$$\neg((M \rightarrow E) \wedge (P \rightarrow E)) \wedge ((M \rightarrow E) \wedge (P \rightarrow E)) \quad (9)$$

Wenn wir die oben eingeführten 4 Aussagen (i) Supervenienz (P2), (ii) kausale Geschlossenheit, (P5), (iii) Nicht-Reduzierbarkeit (P4) und (iv) kausale Exklusion (P5) beibehalten wollen, und wenn das obige Argument schlüssig ist, dann führt die Einführung der zur Debatte stehenden Prämisse der mentalen Verursachung (P1) in einen Widerspruch. Folglich müsste die fragliche Prämisse P1 aufgegeben werden und die These des nicht-reduktiven Physikalismus wäre falsch. Ich meine jedoch, dass es auch möglich ist, die Prämisse P4 aufzugeben, ohne notwendigerweise in einen Widerspruch zu den zentralen Prämissen des nicht-reduktiven Physikalismus zu geraten, und dass deshalb die fragliche Prämisse P1 beibehalten werden kann. Im folgenden Teil werde ich diese Meinung begründen.

### 3 Relationale Eigenschaften und Kausalität des Mentalen

#### 3.1 Mentale Zustände sind an sich relational

##### 3.1.1 Relation zum Subjekt

Eine der Besonderheiten aller mentaler Zustände ist, dass sie ein Subjekt haben [Searle, 2000, S. 561]. In Searles Begrifflichkeit haben mentale Zustände eine erste-Person-Ontologie, das heisst, sie existieren, — als mentale Zustände — nur, wenn es ein Subjekt gibt, das sie erlebt. Das ist in der Struktur der Aussage analog zu einem Satz wie „Omeletten sind mein Lieblingsessen“: Omeletten existieren — als Lieblingsessen — nur, wenn es jemanden gibt, der sie als das bezeichnet. Wenn es niemanden gibt, der Omeletten als ihr Lieblingsessen bezeichnet, ändert das nichts an den Eigenschaften der Omeletten selber, sie sind bloss nicht mehr jemandes Lieblingsessen. Analog verändern sich die physischen Eigenschaften von materiellen Zuständen nicht notwendigerweise in Abhängigkeit davon, ob sie durch ein Subjekt erlebt werden. Was sich aber notwendigerweise damit ändert ist die Eigenschaft, auch ein mentaler Zustand zu sein. Bestimmte materielle Zustände erhalten somit das zusätzliche Attribut, die materielle Basis eines subjektiven mentalen Zustands zu sein. Das Subjekt attribuiert diese Zustände damit als sich eigen [Gallagher, 2000].

##### 3.1.2 Privater Zugang in Form von Qualia

Eng verbunden mit der Subjektivität von mentalen Zuständen ist, dass mentale Zustände sogenannte *Qualia* haben. Es gibt, wie es Thomas Nagel [Nagel, 1974] ausdrückt, für das Subjekt ein „wie es ist, dass“, wenn ein Subjekt einen bestimmten mentalen Zustand hat. Weiter sind diese subjektiven Qualitäten mentaler Zustände *privat*, das heisst, es gibt immer nur genau eine Person, die einen Satz sagen kann wie „Ich habe Schmerzen“, ohne die Bedeutung des Satzes zu ändern: Was den Satz aus der Perspektive des Subjekts, und nur aus der Perspektive des Subjekts wahr macht ist ein unmittelbares Erleben. Im Gegensatz dazu ist das, was einen ähnlichen Satz wie „Er hat Schmerzen“ möglicherweise wahr macht, eine Inferenz aus für Schmerzen typische Verhaltensweisen [Putnam, 1968].

Mentale Zustände zeichnen sich also durch besondere relationale Eigenschaften zum Subjekt aus. Diese Eigenschaften sind, scheint mir, sowohl notwendig (keine mentalen Zustände

ohne Subjekt) wie auch hinreichend (jeder Zustand, der ein Subjekt hat ist auch ein mentaler Zustand) für die Eigenschaft, ein mentaler Zustand zu sein. Zusätzlich beziehen sich diese Eigenschaften auf etwas, für das es in physikalischen Begrifflichkeiten keine Korrespondenzen gibt, etwa Qualia oder Privatheit.

### 3.2 Leiblichkeit als mentale Relationen zum eigenen Körper

Wenn mentale Zustände wie Wünsche auf die materielle Aussenwelt kausal wirksam sein sollen, müssen sie sich als Bewegung des Körpers manifestieren. Diesen Zusammenhang hat Arthur Danto als sogenannte basale Handlungen (basic actions) beschrieben [Danto, 1965]. Basale Handlungen sind für Danto keine Folge des Wollens, sondern ein Akt des Wollens. Wenn ich eine willkürliche Körperbewegung ausführe, ist das für Danto also nicht eine Folge meines Wunsches, mich entsprechend zu bewegen, sondern identisch mit dem Wunsch, mich entsprechend zu bewegen [Danto, 1965, S.148]. Auch in Davidsons klassischer Definition einer Handlung [Davidson, 1963] ist das Vorhandensein eines Wunsches in Kombination mit entsprechender Mittel-Zweck-Überzeugung notwendig und hinreichend für eine Handlung.

Nun möchte ich hier festhalten, dass unser Körper sowohl physische wie auch mentale Eigenschaften hat. Die physischen Eigenschaften sind offensichtlich, ein Körper hat, wie alle anderen materiellen Gegenstände auch, ein Volumen, ein Gewicht, einen Ort, eine Temperatur und verhält sich zumindest mutmasslich genau wie alle anderen materiellen Gegenstände auch ausschliesslich gemäss den Gesetzen der Physik. Gleichzeitig hat der eigene Körper auch mentale Eigenschaften. Er „gehört“ einem Subjekt, das einen qualitativen und privaten Zugang zu den Zuständen des eigenen Körpers hat: Wir erleben unseren Körper in einer Weise, die gänzlich anders ist als das, was eine physische Beschreibung des Zustands unseres Körpers aussagen könnte, er ist vielleicht müde, fühlt sich wohl an oder schmerzt an einer Stelle. Gernot Böhme verwendet für das, was die mentale Beschreibungsebene von körperlichen Zuständen umfasst, die Begriffe des *Leibs* und der *Leiblichkeit* und kontrastiert diese Begriffe mit der physischen Beschreibung des *Körpers* [Böhme, 2008]. Im Folgenden werde ich die Begriffe im obigen Sinn verwenden, um zwischen mentalen und materiellen Beschreibungsweisen des eigenen Leibs beziehungsweise des Körpers zu unterscheiden. Das impliziert jedoch nicht, dass sich Körper und Leib in ihrer Extension unterscheiden müssen.

Analog zu den oben dargestellten Eigenschaften des Mentalen zeichnet sich also auch der Leib durch besondere relationale Eigenschaften zum Subjekt aus. Diese Eigenschaften sind, scheint mir, ebenso notwendig (kein leibliches Erleben ohne Subjekt) wie auch hinreichend (jeder körperliche Zustand, der ein Subjekt hat, ist auch ein leiblicher Zustand) für die Eigenschaft, ein leiblicher Zustand zu sein.

Lange nicht jeder Zustand des eigenen Körpers ist ein leiblicher Zustand (zum Beispiel im Fall einer Anästhesie, oder wenn man den Zustand des Knochenmarks betrachtet) und nicht jeder leibliche Zustand ist ein körperlicher Zustand (zum Beispiel im Fall von Phantomgliedern oder wenn wir träumen). Es gibt aber offensichtlich eine genügend grosse Anzahl von a-posteriori Identitäten von leiblichen und physischen Zuständen, zum Beispiel, dass typischerweise Empfindungen wie Berührtwerden, Warm-/ Kaltsein oder Schmerz in systematischer Weise zusammenhängen mit bestimmten physischen Eigenschaften wie dem Vorhandensein von Druck auf die Haut, dem Fluss von thermischer Energie oder dem Erregungszustand von Nervenfasern eines bestimmten Typs. Und es gibt wohl, wie ich im Abschnitt 6 noch etwas weiter ausführen werde, auch gewisse Normen für das Möglichsein solcher Zusammenhänge.

Solche *Token-Token-Identitäten* von leiblichen und körperlichen Zuständen erklären, dass sich ein leibliches Empfinden einstellt, wenn etwas physisch kausal auf den Körper wirkt. Ebenso erfahrbar ist, dass es auch für Teile derjenigen körperlichen Kausalbeziehungen, die vom Körper auf die Aussenwelt wirken, zumindest Token-Token-Identitäten mit leiblichem Erleben gibt. Das ist etwa der Fall, indem wir erleben, eine körperliche Anstrengung zu leisten, oder indem wir Widerstand erleben, wenn wir Kraft anwenden. Gallagher [2007] bezeichnet dieses leibliche

Empfinden von kausaler Wirkung als „Sense of Agency“<sup>1</sup>.

Die typische Form des leiblichen Erlebens von kausaler Wirkung des Körpers ist wohl, auch im Sinne von Dantos Verständnis von basalen Handlungen, dass ich einen Teil der Bewegungen meines physischen Körpers mit dem Wollen identifiziere, meinen Körper entsprechend zu bewegen. Dieses Wollen hat nun, analog zu allen anderen mentalen Zuständen, wiederum ein Subjekt, dessen Wunsch es ist, dass eine entsprechende Bewegung stattfindet: Es gibt, analog zu den sensorischen Empfindungen, wiederum immer nur genau eine Person, die einen Satz sagen kann wie „Ich will meinen Arm heben“, ohne die Bedeutung des Satzes zu ändern: Was den Satz aus der Perspektive des Subjekts, und nur aus der Perspektive des Subjekts wahr macht ist ein privater und subjektiver mentaler Zustand des Wollens. Im Gegensatz dazu ist das, was einen ähnlichen Satz wie „Sie will ihren Arm heben“ möglicherweise wahr macht, eine Inferenz aus für Willkürhandlungen typischen Äusserungen und Verhaltensweisen.

### 3.3 Relationale Eigenschaften von Handlungen

Wenn sich einfache Körperbewegungen zu Handlungen zusammenfügen, ergeben sich weitere relationale Eigenschaften dieser Körperbewegungen. Diese relationalen Eigenschaften lassen sich anhand des Erklärungsmodells einer Handlung von Davidson gut darstellen. Davidsons Modell erklärt Handlungen durch ihre sogenannte *Rationalisierung* [Davidson, 1963, S. 4]. Solche Rationalisierungen bestehen nach Davidson immer aus zwei Komponenten: Einem *Wunsch* und einer *Überzeugung* [Davidson, 1963, S. 3].

1. Als *Wunsch-Komponente* bezeichnet Davidson eine Pro-Attitüde gegenüber jeglichen Handlungen, die zu einem bestimmten Ereignis führen. Wenn ich zum Beispiel den Wunsch habe, ein Bier zu trinken, ist das für Davidson identisch mit der Aussage, dass ich eine Pro-Attitüde habe zu der Menge aller Handlungen, die dazu führen würden, dass ich Bier trinke.
2. Die *Überzeugungs-Komponente* der Rationalisierung besagt, dass eine bestimmte Handlung in die Menge aller Handlungen fällt, die der Wunsch-Komponente entsprechen, die also ein wunschgemäßes Ereignis zur Folge haben sollten und damit einen Zweck erfüllen. Ich kann zum Beispiel der Überzeugung sein, dass ich Bier trinken werde, wenn ich aufstehe und zum Kühlschrank gehe, weil ich weiss, dass im Kühlschrank Bier ist.

Die beiden Komponenten zusammen bilden das, was Davidson unter *primären Gründen* („primary reasons“) versteht. Ihr Vorhandensein ist im Davidsonschen Modell notwendig und hinreichend für das Vorhandensein einer Handlung, was in der Handlungstheorie als praktischer Syllogismus bezeichnet wird. Ob dieses Modell genügt, um Handlungen vollständig zu erklären, ist umstritten (siehe z. B. [Sehon \[2010\]](#)), für das Argument, das ich ausführen will, genügt es aber, davon ausgehen zu können, dass Wünsche und Überzeugungen notwendiger Teil einer Handlung sind, und dass die Handlung gleichzeitig durch Bewegungen des Körpers realisiert wird.

Wichtig ist mir hier aber die Feststellung, dass sowohl die Wunsch-Komponente wie auch die Überzeugungs-Komponente nun der physischen Realisierung der Handlung weitere, erst einmal in mentalen Begriffen gefasste relationale Eigenschaften attribuieren: Die Wunsch-Komponente enthält solche mental bestimmten relationalen Eigenschaften wie Pro-Einstellung des Subjekts

---

<sup>1</sup>Das Erleben von „aktiver“ Leiblichkeit ist aber typischerweise deutlich schwächer als das die „passive“, Form. Das ist nicht unerwartet, weil vorhersehbare Ereignisse, zu denen die Folgen des eigenen Handelns idealerweise gehören, deutlich weniger Aufmerksamkeit erregen als Ereignisse, die nicht vorhersehbar sind. Deshalb werden sie subjektiv auch weniger intensiv bis sogar gar nicht bewusst wahrgenommen. Gallagher spricht davon, dass dieser primäre Sense of Agency ein recessiver oder marginal Aspekt unseres Bewusstseins ist [[Gallagher, 2012, S. 16](#)]

gegenüber einem durch diese relationalen Eigenschaften definierten Set von physischen Ereignissen. Die Überzeugungs-Komponente enthält als Überzeugungen zu physischen Ereignissen ebenfalls solche relationalen Eigenschaften, wie etwa die, dass sie einen bestimmten Zweck erfüllen können. Und es liessen sich, wenn der Kontext erweitert wird, noch weitere relationale Eigenschaften von Handlungen finden. Handlungen betreffen zum Beispiel häufig weitere Personen, infolge dessen sie weitere relationale Eigenschaften haben. Und einzelne Handlungen sind häufig Teil eines Narrativs, was ihnen weitere relationale Eigenschaften attribuieren kann.

Bestimmte kausal wirksame materielle Zustände erhalten somit das zusätzliche Attribut, die materielle Basis eines subjektiven mentalen Zustands des Wollens zu sein. Das Subjekt attribuiert diese kausal wirksamen Zustände damit ebenfalls als sich eigen.

## 4 Ungleiche relationale Eigenschaften und Identität

In seinem Aufsatz „Über Sinn und Bedeutung“ stellt Frege [1892] fest, dass sich zwei Bezeichnungen desselben Gegenstands in der Art des Gegebenseins unterscheiden können. Die Feststellung, dass unterschiedlich Gegebenes denselben Gegenstand bezeichnet, sei eine oft wertvolle Erweiterung unserer Erkenntnis und *a priori* nicht immer zu begründen [Frege, 1892, S. 25]. Dass „Morgenstern“ und „Abendstern“ denselben Gegenstand bezeichnen, ist, so Frege, eine nicht triviale Erkenntnis, weil uns derselbe Gegenstand als „Morgenstern“ und als „Abendstern“ jeweils unterschiedlich gegeben ist. Frege nennt die Art des Gegebenseins eines Gegenstandes den Sinn der Bezeichnung [Frege, 1892, S. 26]. Sätze, die sich im Sinn unterscheiden, die aber immer dieselben Wahrheitswerte haben, bezeichnen für Frege identische Gegenstände [Frege, 1892, S.36].

Wenn der Wahrheitswert des Vorhandenseins von primären Gründen (also bestimmten Wünschen und Überzeugungen), wie Davidson es annimmt, immer mit dem Wahrheitswert des Vorhandenseins eines der Handlung physisch entsprechenden Zustands übereinstimmt, dann bezeichnen die primären Gründe einerseits, und der physische Zustand des Handelns andererseits, denselben Gegenstand. Gleichzeitig unterscheidet sich aber das Vorhandensein von primären Gründen vom Vorhandensein des physischen Zustand des Handelns in der Art der Gegebenheit: Primäre Gründe sind uns gegeben als Subjekt eines mentalen Zustands, während entsprechende physische Zustände objektiv und unabhängig vom Subjekt feststellbar sein müssen. Die beiden Prämissen des praktischen Syllogismus (das Vorhandensein von als mentale Zustände gegebenen Wünschen und Überzeugungen) unterscheiden sich von der Konklusion (dem Vorhandensein der als physischer Zustand gegebenen entsprechenden Handlung) also, in Freges Begrifflichkeit, in ihrem Sinn, sind aber im Gegenstand, den sie bezeichnen, *a posteriori* identisch.

Wichtig ist mir hier, dass sich die Gegebenheit von primären Gründen von der Gegebenheit entsprechender physisch feststellbarer Zustände nun nicht nur beiläufig, sondern notwendigerweise unterscheidet. Das folgt aus der weiter oben gemachten Feststellung, dass es ein Alleinstellungsmerkmal von mentalen Zuständen ist, Zustände eines Subjekts zu sein. Primäre Gründe sind, in Freges Begrifflichkeit, also notwendigerweise sinnverschieden vom physischen Vorhandensein einer Handlung, weil primäre Gründe notwendigerweise die primären Gründe eines Subjekts sind, während die nicht-mentale Bezeichnung desselben Gegenstands als physischer Zustand nie einen Bezug zu einem Subjekt haben kann. Es gibt also notwendigerweise Unterschiede in den relationalen Eigenschaften von mentalen und physikalistischen Kausalerklärungen, obwohl beide Erklärungen dasselbe zum Gegenstand haben können.

## 5 Folgen für das Argument von Kim

Im obigen Teil habe ich gezeigt, dass mentale Ursachen mit physischen Ursachen *a-posteriori* identisch sein können, sich aber in ihren relationalen Eigenschaften von physischen Ereignissen notwendigerweise unterscheiden. Wenn das richtig ist, dann kann die Prämisse 4 (Nicht-Identität



von mentalen und physischen Ursachen) aus dem Argument von Kim abgelehnt werden. Damit ist das Argument als Ganzes nicht mehr schlüssig. Mentale Ursachen wären damit nicht unabhängig von physischen Ursachen kausal wirksam, sie könnten aber auch nicht auf physische Ursachen reduziert werden, weil es notwendigerweise relationale Eigenschaften gibt, die nur für mentale, nicht aber für physische Erklärungen von Ursachen bestehen. Das ist kompatibel mit der Position des nicht-reduktiven Physikalismus, die Kim mit seinem Argument angreifen will.

### 5.1 Reduktionismus von relationalen Eigenschaften

Aus der Position von Kim ergibt sich daher die Notwendigkeit zu zeigen, dass a-posteriori Identität von mentalen und physischen Ursachen bei gleichzeitig unterschiedlichen relationalen Eigenschaften für sein Argument unproblematisch ist.

Interessanterweise findet sich in Kims Arbeit von 2005 genau in diesem Kontext ein Argument, dass dieses Problem zu antizipieren scheint, das mir aber schwierig scheint [Kim, 2005, S. 34]: Kim definiert, wie ich es hier auch tue, einerseits die Nicht-Reduzierbarkeit des Mentalen durch etwas Physisches als Nicht-Identität von Eigenschaften („Mental properties are not reducible to, and are not identical with, physical properties“). In seinem Argument verweigert er dann aber diesen Eigenschaften, nicht aber dem Gegenstand, auf den sie sich beziehen eine materielle Grundlage: („... all we need to assume here is that physically irreducible properties remain outside the physical domain—that is, if anything is physically reduced, it must be identical with some physical item“). Diese Annahme scheint mir problematisch: Entweder haben die relevanten Eigenschaften per Definition kein entsprechendes physisches Pendant. Dann kann Kims Argument wohl wirklich nichts anderes zeigen, als dass mentale Zustände keine kausale Wirkung haben können. Kim bliebe es dann aber schuldig zu zeigen, warum das Supervenienzprinzip, das für sein Argument so wichtig ist, zwar für kausale Eigenschaften, nicht aber für alle anderen Eigenschaften auch gelten soll.

Es wird also für die Verteidigung des Arguments von Kim notwendig, das reduktionistische Vorgehen auf die relationalen Eigenschaften auszuweiten. Es gilt also nicht nur zu zeigen, dass auch relationale mentale Eigenschaften mit physischen relationalen Eigenschaften a-posteriori identisch sein können, sondern auch, dass diese physisch definierten relationalen Eigenschaften die für Kims Argument anscheinend problematischen mentalen relationalen Eigenschaften vollständig „wegerklären“ können. Damit wäre, analog zu seinem Argument gegen die Möglichkeit von kausalen Eigenschaften von mentalen Zuständen, „keine Luft mehr“ für mentale relationale Eigenschaften: Sie müssten verstanden werden als ein nicht eigentlich relationales, weil zu physisch notwendigen relationalen Eigenschaften redundantes mentales Epiphänomen. In Folge wäre die Unterscheidbarkeit von mentalen und physischen Ursachen anhand substanzialer relationaler Eigenschaften nicht mehr gegeben: Weil alle relationalen Eigenschaften immer wesentlich physisch wären, könnten sie sich nicht zwischen mentalen und physischen Ursachen unterscheiden. Damit müsste entweder die Unterscheidung zwischen mentalen und physischen Ursachen aufgegeben werden, (was der reduktionistischen Agenda entsprechen würde) oder es gäbe wieder keine Erklärung dafür, wie mentale Zustände kausale Wirkung entfalten könnten.

Ich habe diese Argument folgendermassen formalisiert:

- Alle mentalen Ursachen  $M$  haben mentale relationale Eigenschaften  $R_m$  zum Subjekt  $S$ :

$$\forall(M \rightarrow E)\exists R_m(M, S) \quad (10)$$

- Aus der Aussage (i) von oben (Realisierung mentaler Zustände durch physische Zustände) folgt: Die mentalen relationalen Eigenschaften  $R_m$  zwischen  $M$  und  $S$  werden durch physische relationale Eigenschaften  $R_p$  realisiert, die die physische Ursache  $P$  in Bezug setzen zum Körper  $K$  des Subjekts  $S$ :

$$\exists R_p(P, K)(R_m(M, S) \mapsto R_p(P, K)) \quad (11)$$

- Mentale relationale Eigenschaften sind nicht auf physische relationale Eigenschaften reduzierbar:

$$R_m(M, S) \neq R_p(P, K) \quad (12)$$

- Analog zum Exklusionsprinzip aus Aussage (iv) sind relationale Eigenschaften (üblicherweise) nicht überbestimmt. Daraus folgt:

$$R_m(M, S) \neq R_p(P, K) \rightarrow \neg(R_m(M, S) \wedge R_p(P, K)) \quad (13)$$

- Weil aber  $R_p(P, K)$  durch 11 gesetzt ist, folgt aus 13:

$$\therefore (R_m(M, S) = R_p(P, K)) \vee \neg R_m(M, S) \quad (14)$$

Der erste Teil der Disjunktion 14 beinhaltet den reduktionistischen Arm des Arguments. Diesen Teil der Disjunktion zu akzeptieren wäre einerseits für nicht-reduktive Physikalisten wohl grundsätzlich problematisch. Andererseits würde er den Einwand aufheben, dass sich mentale Ursachen von physischen Ursachen in ihren relationalen Eigenschaften unterscheiden können, und dass Kims Argument deshalb nicht schlüssig ist.

Der zweite Teil der Disjunktion 14 drückt die für nicht-reduktive Physikalisten die ebenfalls nicht akzeptable These aus, dass es keine mentalen relationalen Eigenschaften von Ursachen gibt, sondern dass das, was damit gemeint ist, höchstens ein Epiphänomen ist. Auch das wäre für die Verteidigung von Kims Argument natürlich hinreichend.

Ich glaube aber, dass sich dieser Gegeneinwand entkräften lässt. Es gibt dafür zumindest zwei Argumente, die ich im Folgenden zu entwickeln versuche:

### 5.1.1 Zur Überbestimmtheit relationaler Eigenschaften

Es scheint relativ plausibel, die Überbestimmtheit von kausalen Erklärungen im Allgemeinen als problematisch zu betrachten: Sowohl wenn man von kontrafaktischen Theorien der Kausalität ausgeht [Lewis, 1973], wie auch wenn man interventionistische Theorien von Kausalerklärungen [Woodward, 2008] heranzieht, ist es problematisch, Kausalität feststellen zu wollen, wenn für eine kausal zu erklärende Folge  $E$  mehrere genügende Ursachen ( $P_1, P_2, P_3, \dots$ ) gleichzeitig vorhanden sind. Das gleichzeitige Vorhandensein mehrerer genügender Ursachen für dieselbe Folge verletzt die epistemischen Rahmenbedingungen, unter denen kausale Abhängigkeit eindeutig festgestellt werden können. Ein klassisches Beispiel für eine solche problematische kausale Überbestimmtheit ist der Fall, wo eine Fensterscheibe kaputt geht, nachdem sie gleichzeitig von zwei Steinen getroffen wurde, von denen jeder allein in der Lage war, die Scheibe einzuschlagen. Es ist unklar, wie wir eindeutig feststellen sollen, ob je einer der beiden Steine verursacht hat, dass die Scheibe zerbricht, weil sie ja immer auch zerbrechen würde, wenn einer der beiden Steine nicht treffen würde.

Weil die Epistemik von relationalen Eigenschaften aber ganz anders ist als die von kausalen Erklärungen, ist es nicht klar, warum eine Überbestimmtheit von relationalen Eigenschaften problematisch sein sollte. Im Gegensatz zu kausalen Beziehungen bleiben nämlich alle relationale Eigenschaften eindeutig feststellbar, auch wenn ein Subset dieser Eigenschaften hinreichend wäre, bestehende Relationen eindeutig zu beschreiben. Und im Gegensatz zu so exotischen Fällen wie der Fensterscheibe, die gleichzeitig von zwei Steinen getroffen wird, sind solche Überbestimmtheiten von relationalen Eigenschaften sehr alltäglich. Ein Beispiel:

„Zu meinem Verhältnis zu Frau M lässt sich sagen, dass ich der Sohn von Frau M bin, und dass Frau M meine Mutter ist. Zu meinem Verhältnis zu Herrn V lässt sich sagen, dass ich ebenfalls der Sohn von Herrn V bin, und dass Herr V mein Vater ist. Zum Verhältnis von Frau M und Herrn V lässt sich sagen, dass sie gemeinsam meine Eltern sind.“

Im obigen Beispiel gibt es viele Beziehungen, die nicht identisch sind, die sich aber gegenseitig bedingen. Mein Verhältnis zu Frau M als ihr Sohn ist zum Beispiel ein anderes Verhältnis als ihr Verhältnis zu mir als meine Mutter, auch wenn das eine Verhältnis das andere jeweils notwendigerweise nach sich zieht. Und die zwei Beziehungen vom Typ „Eltern-Kind“ definieren eine dritte Beziehung vom Typ „gemeinsame Eltern“ zwischen den zwei Elternteilen vollständig. Es würde aber wohl niemandem einfallen, aus dieser Überbestimmtheit folgern zu dürfen, dass eine dieser Beziehungen eliminiert werden müsste oder ein Epiphänomen ist. Auch wenn Teile oder die Gesamtheit solcher Beziehungen auf verschiedenen Ebenen definiert sind, würde man eine Überdeterminiertheit von relationalen Eigenschaften kaum als problematisch betrachten: Gemeinsam meine leiblichen Eltern zu sein bleibt eine gehaltvolle (das heisst nicht epiphänomenale) Aussage, auch wenn sie sich materiell vollständig dadurch ergibt, dass es einen Organismus gibt, der aus der Vereinigung der Gameten der beiden Eltern entstanden ist.

Eine analoge Sachlage lässt sich nun auch für leibliches Erleben von Kausalität (also Galaghers Sense of Agency) konstruieren. Dass sich mir als Subjekt kausale Wirksamkeit durch einen besonderen leiblichen und damit mentalen Bezug zu einem Teil der Welt (nämlich dem, der meinem Körper entspricht) ergibt, mag gleichzeitig vollständig durch materiell definierte Beziehungen zwischen den Teilen dieses Körpers bestimmt sein. Daraus folgt aber wieder nicht, dass die leibliche Darstellung der relevanten Bezüge eliminiert werden müsste.

Im Gegensatz zu Kausalbeziehungen ist die Überbestimmtheit von relationalen Eigenschaften also unproblematisch, denn es gibt keine epistemische Notwendigkeit, „überschüssige“ relationale Eigenschaften zu eliminieren. Im Gegenteil, eine solche Elimination wäre falsch, weil sie sinnvolle Aussagen wie „Frau M und Herr V sind meine Eltern“ oder „Ich habe das Fenster aufgemacht“ ungültig machen würde. Damit wird die Prämisse 13 falsch, und das Argument verliert seine Schlüssigkeit.

### 5.1.2 Die Erklärungslücke

Ein weiteres Problem für eine reduktionistische Agenda von leiblich erlebten relationalen Eigenschaften (wie sie mentale Kausalerklärungen notwendigerweise haben) ist, dass sich diese relationalen Eigenschaften in Art und Weise manifestieren, die es völlig unklar macht, wie diese reduktionistische Agenda denn überzeugend gelingen könnte. Mentale relationale Eigenschaften von Kausalerklärungen haben, wie andere mentale Zustände auch, ein Subjekt (siehe Abschnitt 3.1.1) und Qualia (siehe Abschnitt 3.1.2). Das heisst, sie beinhalten notwendigerweise ein Subjekt, das ein Ereignis erklärt als „weil ich“ und sie manifestieren sich für dieses Subjekt in einer gänzlich unphysikalischen Weise als zum Beispiel anstrengend, lustvoll oder angenehm. Es ist aber grundsätzlich unklar, wie eine reduktionistische Erklärung von solchen Arten von Eigenschaften in der Begrifflichkeit von rein physischen Eigenschaften überhaupt ansetzen könnte [Chalmers, 1995]. Das gilt auch für die relationalen Eigenschaften von mentalen Kausalerklärungen. Sie sind zwar ausserhalb der Erklärungsmöglichkeiten, die die Biologie oder Physik anbieten kann, damit aber nicht notwendigerweise ausserhalb der Domäne des Physischen.

Mit dem Hinweis auf die besonderen Qualitäten der relationalen Eigenschaften von mentalen Kausalerklärungen erweitert sich der zu erklärende Gegenstand also in einer Weise, die deutlich und in einer problematischen Weise über das hinausgeht, was Kims Argument beinhaltet: Einerseits scheint es nicht zulässig, mentale Verursachung ohne diese relationalen Eigenschaften verstehen zu wollen, denn sie sind essenziell durch diese relationalen Eigenschaften mental. Umgekehrt müssten mit dem Einschluss solcher relationalen Eigenschaften in das Argument für all die nur zu gut bekannten allgemeineren Probleme von reduktionistischen Erklärungen von mentalen Zuständen befriedigende Antworten möglich sein. Das scheint aber aus tiefgehenden Gründen problematisch zu sein. Die Frage nach der Möglichkeit kausaler Wirksamkeit von mentalen Zuständen lässt sich also nicht ganz trennen von den anderen Problemen, die sich stellen, wenn mentale Zustände auf physische Zustände reduziert werden sollen. Etwas schärfer formu-

liert: Ein Problem von Kims Argument scheint zu sein, dass es das, was kausale Wirksamkeit von mentalen Zuständen umfasst erst in einer wie es sich zeigt unzulässigen, weil vom Subjekt abstrahierten Weise vereinfacht, um dann zu zeigen, dass dieses vereinfachte Verständnis von mentalen Ursachen nicht in der Lage ist, das zu leisten, was es sollte, nämlich mit einem nicht-reduktiven Physikalismus kompatibel zu sein.

## 6 Normative Relevanz von angemessenen kausalen Selbstattributionen

Ich will zum Schluss dieses Essays noch kurz darauf eingehen, dass kausale Selbstattributionen, wie sie die oben besprochenen mentalen Zustände beinhalten, nicht nur notwendig sind, weil sie einen Zustand zu einem mentalen Zustand machen. Sie sind in der Begrifflichkeit von Shaun Gallagher der Kristallisationskern dessen, was er als „minimales Selbst“ bezeichnet: Ein immediates Bewusstsein meiner Selbst als Subjekt von Erfahrungen. Diese Erfahrungen beinhalten einen „Sense of Ownership“, das heisst die Erfahrung, dass ich es bin, die/der etwas erfährt, und, für die gegebene Fragestellung zentral, einen „Sense of Agency“, das heisst die Erfahrung, dass ich es bin, die/der etwas verursacht [Gallagher, 2000]. Ungewöhnliche Selbstattributionen führen damit zu ungewöhnlichem Selbsterleben.

In Tabelle 1 sind vier Sätze mit unterschiedlichen Kausalattributionen. Diese Attributionen unterscheiden sich in der vertikalen Achse im Vorhandensein einer kausalen Selbstattribution, in der horizontalen Achse darin, ob mutmasslicherweise ein leiblicher Bezug zur gegebenen Ursache angenommen werden kann. Aus der Perspektive des physikalischen Reduktionismus allein sind die kausalen Erklärungen des Ereignisses, dass das Fenster aufgeht, in Satz 1 und 2 vom selben Typ, während weder Satz 3 noch Satz 4 eine valide kausale Erklärungen sind. Das hat eine Reihe von Folgen, die mir problematisch scheinen. Ich werde im Folgenden kurz darlegen warum.

Tabelle. 1: Abhängigkeit von leiblichem Bezug zur physischen Ursache und zur kausalen Selbstattribution.

	Leiblicher Bezug zur physischen Ursache unplausibel	Leiblicher Bezug zur physischen Ursache plausibel
Keine kausale Selbstattribution (Wollen nicht notwendig)	„Als Folge komplexer physischer Vorgänge weht der Wind so, dass das Fenster aufgeht.“ <b>(Satz 1)</b>	„Als Folge komplexer physischer Vorgänge bewegt sich mein Körper so, dass das Fenster aufgeht.“ <b>(Satz 2)</b>
Kausale Selbstattribution (Wollen notwendig)	„Als Folge meines Wunsches nach Abkühlung weht der Wind so, dass das Fenster aufgeht.“ <b>(Satz 3)</b>	„Als Folge meines Wunsches nach Abkühlung bewegt sich mein Körper so, dass das Fenster aufgeht.“ <b>(Satz 4)</b>

Ein erstes Problem ergibt sich, weil Satz 2 zwar möglicherweise richtig, aber im Gegensatz zu Satz 1 in einer relevanten Weise unvollständig ist: Er würde zulassen, dass sich mein Körper in der beschriebenen Weise bewegt, ohne dass ich das will. Das wäre, im Gegensatz zu dem, was Satz 1 beschreibt, eine sehr befremdliche Erfahrung. In der Psychiatrie spricht man in solchen Fällen von einem Depersonalisierungserleben, und geht davon aus, dass Personen, die solche Erfahrungen regelmässig machen, typischerweise psychisch krank sind und behandelt werden sollten.

Das zweite Problem stellt sich, weil unter den Prämissen des physikalischen Reduktionismus unklar ist, warum Satz 3 im Gegensatz zu Satz 4 so bizarr wirkt, wenn sich Satz 3 und 4

in ihrer kausalen Erklärungskraft nicht unterscheiden: Wenn Satz 4 zwar plausibel scheint, es aber a-priori unmöglich ist, dass der mentale Zustand des Wünschens ein physisches Ereignis verursachen kann, sollte Satz 3 doch ebenso (un-)plausibel sein wie Satz 4. Wiederum sind in den Augen der Psychiatrie Personen, die wiederholt von Erfahrungen in der Art der Aussage von Satz 3 berichten typischerweise psychisch krank: Sie leiden unter etwas, das man als wahnhafte Ego-Ausweitung bezeichnet und das ebenfalls behandelt gehört. Weil in Kims Position Satz 3 und Satz 4 in gleicher Weise falsch sind (Ein mentaler Zustand wie der des Wünschens kann nie eine kausale Wirkung haben), würden nun auch Personen, die regelmässig Aussagen von der Art von Satz 4 machen an einer wahnhaften Ego-Ausweitung leiden. Sie wären folglich gleich „irr“ wie Personen, die Dinge wie in Satz 3 behaupten.

Falls die Psychiatrie also nicht völlig fehl darin geht, wenn sie dem Erleben von der Art wie in den Sätzen 2 und 3 einen Krankheitswert zuschreiben, während Sätze in der Art von Satz 1 und 4 „normal“ sind, muss es Normen geben, die bestimmte Verbindungen zwischen Wollen, leiblichem Erleben und kausaler Wirkung zulassen, und andere nicht. Ich meine, ohne das hier noch weiter ausführen zu können, dass diese Normen eine gewisse Berechtigung haben, und dass wir sie brauchen. Gleichzeitig stehen sie im Gegensatz zu den Konklusionen des physikalischen Reduktionismus. Deshalb scheint mir die Position der physikalischen Reduktionisten in einer relevanten Form unvollständig.

## 7 Diskussion

Im obigen Text habe ich argumentiert, dass sich eine Identität von physischen und mentalen Ursachen auch aus einem nicht-reduktionistischen Verständnis von mentalen Zuständen vertreten lässt, wenn man davon ausgeht, dass es für mentale Zustände spezifische relationale Eigenschaften gibt, die durch den Hinweis auf ihre materielle Realisation nicht einfach eliminierbar sind. Das ist so, weil im Gegensatz zu kausalen Eigenschaften Überbestimmtheit bei relationalen Eigenschaften unproblematisch ist, und weil diese relationalen Eigenschaften sich in Qualitäten manifestieren, für die es keine brauchbaren physikalischen Erklärungen gibt. Mentale Kausalerklärungen bezeichnen dann etwas, dass sich immer auch als physikalische Kausalerklärung bezeichnen lässt und wären deshalb tatsächlich potentiell kausal wirksam. Die Einschränkung, den diese Rechtfertigung der Möglichkeit von mentaler Verursachung für mentale Ursachen mit sich zieht, ist aber beträchtlich: Daraus folgt nämlich gleichzeitig, dass mentale Verursachung nur innerhalb dessen möglich wäre, was die physische Welt für sich allein auch tatsächlich zuliesse.

Dem nachzugehen, ist ein anderes Projekt. Um die obigen Überlegungen aber noch abzuschliessen, möchte ich zum Schluss noch kurz auf die Frage eingehen, ob dieses Verständnis von mentaler Verursachung nicht etwas zu wenig Gehalt hat für das gesamte Phänomen. Ich denke, diese Sorge ist unberechtigt. Die relevanten relationalen Eigenschaften sind nämlich eine notwendige Bedingung für das Vorhandensein von mentalen Zuständen überhaupt, und damit, im Gegensatz zu klassischen Beispielen wie Freges „Morgenstern“ und „Abendstern“ nicht vereinzelte, sondern sehr systematisch bestehende und dauernd vorhandene relationale Eigenschaften. Ohne diese relationalen Eigenschaften gäbe es kein Subjekt. Deshalb äussert sich ihr Fehlen oder Fehl-am-Platz-Sein auch sehr zuverlässig als psychopathologisch relevantes Missempfinden des eigenen Selbst. Für einen nicht-reduktiven Physikalismus ist es kohärent, wenn sich diese relationalen Eigenschaften einerseits als ein erforschbarer neurobiologischer Zustand [Searle, 2000], andererseits als ein subjektiv erlebter „Sense of Agency“ [Gallagher, 2000] feststellbar sind.

## Literaturverzeichnis

L. R. Baker. Metaphysics and mental causation. *Mental causation*, pages 75–95, 1993.

- K. Bennett. Why the exclusion problem seems intractable, and how, just maybe, to tract it. *Noûs*, 37(3):471–497, 2003.
- G. Böhme. *Ethik leiblicher Existenz: Über unseren Umgang mit der eigenen Natur*. Suhrkamp, 2008.
- D. J. Chalmers. Facing up to the problem of consciousness. *Journal of consciousness studies*, 2(3):200–219, 1995.
- A. C. Danto. Basic actions. *American Philosophical Quarterly*, 2(2):141–148, 1965.
- D. Davidson. Actions, Reasons and Causes. *The Journal of Philosophy*, 60:685–700, 1963.
- G. Frege. Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, pages 25–50, 1892.
- S. Gallagher. Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Science*, 4(1):14–21, Jan 2000.
- S. Gallagher. The natural philosophy of agency. *Philosophy Compass*, 2(2):347–357, 2007.
- S. Gallagher. Multiple aspects in the sense of agency. *New Ideas in Psychology*, 30(1):15–31, 2012.
- T. Horgan. Mental Quausation. *Philosophical Perspectives*, 3:47–76, 1989.
- J. Kim. The Supervenience Argument Motivated, Clarified and Defended. In *Physicalism, or Something Near Enough*. Princeton University Press, 2005.
- T. Kroedel and M. Schulz. Grounding mental causation. *Synthese*, 193(6):1909–1923, 2016.
- D. Lewis. Causation. *The Journal of Philosophy*, 70(17):556–567, 1973.
- T. Nagel. What is it like to be a bat? *Philosophical Review*, 83:435–450, 1974.
- H. Putnam. Brains and Behaviour. *Analytical Philosophy*, 11:24–36, 1968.
- J. Searle. Consciousness. *Annual Reviews of Neuroscience*, 23:557–578, 2000.
- S. Sehon. *Teleological Explanation*, chapter 16, pages 121–128. John Wiley and Sons, Ltd, 2010.
- J. Woodward. Mental Causation and Neural Mechanisms. In J. Hohwy and J. Kallestrup, editors, *Being Reduced: New Essays on Reduction, Explanation, and Causation*, pages 218–262. Oxford University Press, Oxford, 2008.