

Der Sitz des Geistes und das Unbewusste – Philosophische Probleme im Lichte Situierter Ansätze

Beate Krickel

1 Einleitung: Das „neue Unbewusste“

Das Unbewusste ist wieder in aller Munde: Populärwissenschaftliche Bücher, die den unbewussten Geist thematisieren, werden internationale Bestseller¹ und es gibt sicherlich keine populärwissenschaftliche Psychologie-Zeitschrift, die nicht schon mehrfach über das Thema berichtet hat². Dieses neue Interesse am unbewussten Geist ist vor allem neuen Erkenntnissen aus der kognitiven Psychologie und der Sozialpsychologie zu verdanken.

In der kognitiven Psychologie zum Beispiel untersuchen Wissenschaftler:innen das Phänomen der *Blindsight*, also das unbewusste Sehen. Blindsight wurde zuerst bei Patient:innen beobachtet, die eine Verletzung des primären visuellen Kortex hatten (Weiskrantz, 1995). Diese Patienten hatten keine bewussten Erlebnisse von visuellen Eindrücken. Trotzdem waren sie in der Lage, akkurat auf Stimuli zu reagieren, wenn sie gebeten wurden, zu raten, wo sich welcher Stimulus befindet.

Ähnliche Beobachtungen konnten auch bei gesunden Versuchspersonen gemacht werden, zum Beispiel mithilfe von Experimenten basierend auf der so genannten *Binokularen Rivalität*. Dabei werden Versuchspersonen gleichzeitig zwei Bilder gezeigt – und zwar so, dass jedes Bild nur von jeweils einem Auge gesehen wird. Dies führt dazu, dass die Bilder sich im Bewusstsein abwechseln. Zeigt man dem einen Auge nun ein Bild, das aus vielen bunten Pixeln besteht (auch „Mondrian“ genannt nach dem niederländischen Maler Piet Mondrian), so bleibt dieses minutenlang bewusst und unterdrückt die bewusste Wahrnehmung des anderen Bildes. In einer Studie konnte gezeigt werden, dass die Versuchspersonen das unterdrückte Bild unter bestimmten Umständen jedoch scheinbar unbewusst wahrnehmen. Zeigt das unterdrückte Bild zum Beispiel eine nackte Person, so wird die Aufmerksamkeit der Versuchsperson von dem Bild je nach sexueller Ausrichtung angezogen oder abgestoßen (Jiang, Costello, Fang, Huang & He, 2006).

Ein Phänomen, das in der Sozialpsychologie untersucht wird und bereits einen großen Einfluss auf öffentliche Debatten zu Rassismus und Sexismus genommen hat, sind so genannte *implicit bias* – scheinbar unbewusste Präferenzen oder Vorurteile, die unser Verhalten

¹ Zum Beispiel: *Thinking, fast and slow* (Kahneman, 2011), *Subliminal: How Your Unconscious Mind Rules Your Behavior* (Mlodinow, 2013), und *Before You Know It: The Unconscious Reasons We Do What We Do* (Bargh, 2017).

² Zum Beispiel: Ayan (2018).

beeinflussen, ohne dass wir dies bemerken (Banaji, Bhaskar & Brownstein, 2015).³ Die Existenz dieser unbewussten Vorurteile wurde abgeleitet aus Beobachtungen wie der folgenden: In Studien konnte gezeigt werden, dass Versuchspersonen, die sich selbst als vorurteilsfrei beschreiben, trotzdem mentale Assoziationen⁴ aufweisen, die dieser Selbstbeschreibung zu widersprechen scheinen. Zum Beispiel zeigen viele Personen, die zum Beispiel Männer nicht für prinzipiell intelligenter halten als Frauen in so genannten *impliziten Testverfahren*, dass sie Männer stärker mit Intelligenz assoziieren als Frauen. Das prominenteste dieser Testverfahren ist der so genannte *implizite Assoziationstest* (kurz: IAT) (Greenwald, McGhee & Schwartz, 1998), der auch online von jedem ausprobiert werden kann.⁵ Weitere scheinbar unbewusste geistige Phänomene, die gegenwärtig von Psycholog:innen untersucht werden, sind unbewusstes Lernen, unbewusstes Entscheidungentreffen, unbewusste Emotionen und unbewusste Gedächtnisinhalte (Hassin, Uleman & Bargh, 2005).

Diese neuen psychologischen Erkenntnisse werden oft als das „neue Unbewusste“ bezeichnet (Hassin et al., 2005). „Neu“ deshalb, da die Phänomene wenig bis gar nichts mit dem „alten“, das heißt psychoanalytischen Unbewussten zu tun haben (aber siehe z. B.: Berlin (2011)). Während das psychoanalytische Unbewusste (auch „dynamisches Unbewusstes“ genannt) durch Verdrängung in Schach gehalten werden musste, um das Ego vor konfliktbeladenen Wünschen und Phantasien zu schützen, scheint das neue Unbewusste eher nüchtern und unaufgeregt – aber nicht weniger faszinierend.

Beim neuen Unbewussten handelt es sich um paradigmatische Beispiele geistiger Phänomene (Sehen, Urteilen, Entscheiden etc.), die *normalerweise* bewusst ablaufen oder von denen wir *bisher davon ausgingen*, dass sie immer bewusst ablaufen (Hassin, 2013). Das Faszinierende am neuen Unbewussten scheint also zu sein, dass Phänomene, von denen wir dachten, dass sie immer bewusst ablaufen, tatsächlich auch unbewusst ablaufen können.

Zentrale Fragen, die sich aus philosophischer Sicht stellen, sind: Inwiefern sind die scheinbar unbewussten Prozesse die gleichen Prozesse wie die, die „normalerweise“ bewusst ablaufen?

³ Es ist wird momentan kontrovers diskutiert, ob unbewusste Vorurteile tatsächlich unbewusst sind. In einer Studie konnte zum Beispiel gezeigt werden, dass Versuchspersonen in der Lage sind, über ihre Assoziationen zu berichten, wenn sie durch das experimentelle Setup dazu gebracht werden, ihre Aufmerksamkeit darauf zu lenken (Hahn, Judd, Hirsh & Blair, 2014). Ich argumentiere dafür, dass dieser Befund nicht zeigt, dass es sich um ein bewusstes Phänomen handelt und das eine Klärung des Begriffs UNBEWUSST notwendig ist (Krickel, 2018).

⁴ Als „mentale Assoziationen“ werden in der Psychologie Verknüpfungen zwischen mentalen Elementen wie Begriffen bezeichnet. Ein einfaches Beispiel: Wenn jemand „Salz“ sagt, denken Sie sicherlich direkt an „Pfeffer“ – Sie assoziieren Salz mit Pfeffer. Diese Assoziation kommt nicht durch eine logische Beziehung zwischen Salz und Pfeffer zustande, sondern unter anderem dadurch, dass bei uns in der Regel Salz- und Pfefferstreuer gemeinsam auf Restauranttischen stehen. Diese Assoziation hat wahrscheinlich Auswirkungen auf Ihr Verhalten: Wenn Sie zum Beispiel einen Pfefferstreuer auf einen Tisch stehen sehen, erwarten Sie, dass dort auch ein Salzstreuer stehen wird, weshalb Sie auf Ihrer Suche nach einem Salzstreuer zu diesem Tisch gehen werden.

⁵ <https://implicit.harvard.edu/implicit/germany/>

Ist unbewussten Sehen tatsächlich ein Fall von *Sehen*? Sind unbewusste Vorurteile und Präferenzen tatsächlich die gleiche Art von Phänomenen, die wir als „bewusste Vorurteile“⁶ bzw. „Präferenzen“ bezeichnen? Allgemein gefragt: Was macht diese unbewussten Phänomene zu *geistigen* Phänomenen eines bestimmten Typs? Warum sollten wir den Schluss ziehen, dass unser *Geist* unbewusste Anteile hat? Ist es nicht offensichtlich, dass viele Einflüsse auf unser Verhalten und Denken uns nicht bewusst sind? Warum also sollte das „neue Unbewusste“ irgendwem überraschen oder verunsichern?

Eine Überlegung, die die Unterscheidung zwischen dem unbewussten Geist und dem, was ich das *bloß Nicht-Bewusste* nennen möchte, motiviert, ist die folgende: Es ist unbestritten, dass in unserem Gehirn, unserem Körper und unserer Umgebung Prozesse ablaufen, die unser Verhalten beeinflussen, denen wir uns nicht bewusst sind. Zum Beispiel passen wir in der Regel automatisch unsere Körperposition an, wenn wir in einem fahrenden Bus stehen, um nicht umzufallen. Ein weiteres Beispiel für Einflüsse auf unser Verhalten, denen wir uns in der Regel nicht bewusst sind, ist das Phänomen, dass im Englischen mit „hanger“ (eine Kombination aus „hunger“ und „anger“) bezeichnet wird. Bei Hunger neigen viele Menschen dazu, einen Streit anzufangen, ohne ihre Streitlust mit ihrem Hunger in Verbindung zu bringen (Bushman, Dewall, Pond & Hanus, 2014). Weitere klare Fälle von Einflüssen, die uns nicht bewusst sind: die Vielzahl von neuronalen Prozessen, die in die Weiterleitung und Verarbeitung von Reizen und Information involviert sind; hormonelle Schwankungen, die unsere Stimmungen und Gedanken färben; Prägungen durch Erziehung oder Gruppenzwang, die unsere Meinungen beeinflussen.

Fälle wie diese sind nicht gemeint, wenn vom „neuen Unbewussten“ die Rede ist. Erstens würde man damit sicherlich keinen Hund hinterm Ofen hervorlocken, da die Existenz solcher Einflüsse wohl niemanden überrascht. Zweitens handelt es sich bei den Beispielen nicht um Prozesse, die vergleichbar wären mit klassischen geistigen Prozessen wie Denken, Entscheidungtreffen, Wahrnehmen und Urteilen (auch wenn die Phänomene sicherlich Einfluss auf diese geistigen Prozesse nehmen). Was also unterscheidet die vermeintlich unbewussten *geistigen* Prozesse von diesen Fällen bloß nicht-bewusster Einflüsse?

Eine Theorie des Geistes, die den Anspruch erhebt, ein vollständiges Bild des Geistes zu liefern, muss in der Lage sein, die gestellte Frage zu beantworten. Falls sie dies nicht tut, verliert sie ihren Gegenstand: *Wovon* ist diese Theorie dann eine Theorie? Die Herausforderung stellt

⁶ Hier gilt zu beachten, dass „Vorurteil“ eine eher unglückliche Übersetzung des englischen Ausdrucks „bias“ ist. In der deutschsprachigen Psychologie wird „bias“ oft mit „Verzerrung“ übersetzt, was aber keine einfache alltagssprachliche Übersetzung hat. In der populärwissenschaftlichen Presse wird daher häufig die Übersetzung „unbewusstes Vorurteil“ verwendet.

sich philosophischen genauso wie wissenschaftlich-psychologischen Theorien jedoch auf unterschiedliche Weise. Philosoph:innen müssen beantworten, was das Kennzeichen des Mentalen (engl.: *mark of the mental*, siehe z.B. Pernu (2017)) ist im Angesicht der wissenschaftlichen Evidenz für einen unbewussten Geist – Bewusstsein wird damit klarerweise als das Kennzeichen des Mentalen ausgeschlossen. Psycholog:innen müssen Bedingungen formulieren, unter denen Daten tatsächlich Evidenz für einen unbewussten *geistigen* Prozess darstellen und nicht bloß Evidenz für einen nicht bewussten, nicht-geistigen (ähnlich der unkontroversen Fälle). Offensichtlich können Philosoph:innen und Psycholog:innen hier wunderbar zusammenarbeiten.

In diesem Aufsatz betrachte ich die Herausforderung aus der Perspektive eines bestimmten neuen Forschungsprogramms – dem so genannten *situierten Geist*. Dieses Forschungsprogramm versammelt eine Vielzahl von unterschiedlichen Theorieansätzen, die die Verkörperung und physische sowie soziale Bedingtheit des Geistigen betonen. Die verschiedenen Ansätze unterscheiden sich teilweise stark. Zu Darstellungszwecken werde ich deren Unterschiede hier ignorieren und nur dann darauf hinweisen, wenn es für meine Argumentation wichtig ist. Unter das Forschungsprogramm des situierten Geistes fasse ich hier Ansätze, die mindestens eine der zwei Kernthesen der klassischen Kognitionswissenschaft ablehnen: die These des *Internalismus* (Geist ist nur im Gehirn) und die des *Computationalismus* (der Geist/das Gehirn ist ein digitaler Computer). Ich werde zeigen, dass der unbewusste Geist eine Herausforderung für Ansätze des situierten Geistes darstellt – im Gegensatz zur klassischen Kognitionswissenschaft, die eine klare Antwort parat hat.

Im Folgenden werde ich zunächst die klassische Sicht auf den Geist erläutern. Dann werde ich erklären, wodurch das Forschungsprogramm des situierten Geistes motiviert ist und dessen Kernunterschiede zum klassischen Ansatz herausstellen. Dann werde ich untersuchen, ob und wie der unbewusste Geist in das klassische Bild und das neue Bild des Geistes passt. Es wird sich zeigen, dass der klassische Ansatz eine einfache Antwort auf unsere Herausforderung hat, die dem situierten Geist nicht zur Verfügung steht. Wie also kann das Forschungsprogramm des situierten Geistes die Unterscheidung von unbewusstem und bewusstem Geist einfangen? Abschließend möchte ich eine Idee skizzieren, wie auch das situierte Forschungsprogramm die Unterscheidung zwischen unbewusstem Geist und dem bloß Nicht-Bewussten bewerkstelligen kann.

2 Die Theorie des Geistes der klassischen Kognitionswissenschaft

Die Kernidee der klassischen Kognitionswissenschaft ist, dass der Geist bzw. das Gehirn funktioniert wie ein (digitaler) Computer: Wie eine Computer-Hardware ein Programm

realisiert, so realisiert das Gehirn den Geist. Ein digitaler Computer realisiert ein Programm, indem bestimmte physikalische Zustände nach bestimmten Regeln so manipuliert werden, dass sie einen bestimmten bedeutungstragenden Input adäquat in einen bedeutungstragenden Output umwandeln können. Ein Beispiel: Über meine Computertastatur und meine Maus kann ich einen bedeutungstragenden Input generieren – zum Beispiel kann ich in meinen Wörterbuch-Programm das Wort „Geist“ eingeben und mit dem Cursor meiner Maus auf „Übersetzen“ klicken. Der Input hat also die Bedeutung „Übersetze das Wort ‚Geist‘ ins Englische“. Dieser Input wird in ein bestimmtes Muster von Strom-an/Strom-aus-Zuständen im Prozessor meines Computers übersetzt. Das Wörterbuch-Programm sorgt nun dafür, dass diese Strom-an/Strom-aus-Zustände nach den Regeln des Programms so manipuliert werden, dass auf meinem Bildschirm ein bedeutungstragender Output in Form verschiedener Pixelmuster erscheint. Zum Beispiel erscheinen dort die Pixelmuster „mind“, „spirit“ und „brandy“. Der Output trägt also die Bedeutung „Das deutsche Wort ‚Geist‘ kann mit den englischen Wörtern ‚mind‘, ‚spirit‘ und ‚brandy‘ übersetzt werden“. Die Regeln des Programms haben die physikalischen Zustände des Computers in geeigneter Weise manipuliert, das heißt so, dass eine korrekte Übersetzung gemacht wurde. Auf die gleiche Art können Computer komplizierte mathematische Berechnungen durchführen, Schach spielen oder uns den Weg zum Bahnhof zeigen. Der Kern der *Computation* ist das regelbasierte Manipulieren von physikalischen Zuständen (Strom an/Strom aus, Pixel etc.), die eine Bedeutung tragen.

Seit dem Ende des Behaviorismus war die Kognitionswissenschaft geprägt von den erstaunlichen Entwicklungen im Bereich der Computerwissenschaft und künstlichen Intelligenz. Diese machten deutlich, wie komplexe logische Aufgaben automatisiert werden konnten, von denen man bis dato angenommen hatte, dass diese sich nicht in physischen Systemen ohne die Steuerung durch einen Menschen realisieren ließen. Der Schluss lag nun nahe, dass man darin nicht nur praktische Werkzeuge gefunden hatte, sondern auch ein Modell davon, wie Menschen solche logischen Aufgaben lösen können – der *Computationalismus* war geboren (auch „repräsentationale Theorie des Geistes“ genannt). Damit war auch eines der größten Rätsel der Philosophie und der Psychologie gelöst: Wie erklärt man die Rationalität (oder auch die *Gründe-Sensitivität*) des Menschen? Philosophen wie René Descartes nahmen an, dass dies nur durch eine immaterielle Seele erklärt werden könne. Nun hatte man eine einfache Erklärung, die mit dem allgemeinen physikalistischen Weltbild vereinbar war: Geistige Prozesse bestehen darin, dass bedeutungstragende neuronale Zustände regelbasiert manipuliert werden. Die Regeln entsprechen den logischen bzw. rationalen Prinzipien, die das menschliche Denken und Handeln ausmachen. Diese werden im Gehirn realisiert, indem die

kausalen Verknüpfungen zwischen den realisierenden physischen Zuständen in unserem Gehirn den rationalen Verknüpfungen den Inhalten der Zustände entsprechen.

Allerdings hat der Computationalismus ein zentrales Problem: Wie kommen die neuronalen Zustände, die regelbasiert manipuliert werden, zu ihrer Bedeutung? Für den digitalen Computer lässt sich diese Frage einfach beantworten: Der Input hat die Bedeutung „Übersetze das Wort ‚Geist‘ ins Englische“, da sich die Entwickler:innen des Computers darauf geeinigt haben und Benutzer:innen diese Konvention lernen und übernehmen. Aus dem gleichen Grund steht das Auftauchen von Pixelmustern auf meinem Bildschirm für „Das deutsche Wort ‚Geist‘ kann mit den englischen Wörtern ‚mind‘, ‚spirit‘ und ‚brandy‘ übersetzt werden“. Und aus dem gleichen Grund werden Strom-an/Strom-aus-Muster mit Sequenzen von Einsen und Nullen beschrieben („Strom an“ bedeutet „1“, „Strom aus“ bedeutet „0“). Mit anderen Worten: Die physikalischen Zustände meines Computers tragen Bedeutung *qua Konvention*. Sie bedeuten etwas, da es Menschen gibt, die die physikalischen Zustände als bedeutungstragend *interpretieren*. Man spricht hier auch von *abgeleiteter Bedeutung*, da sich diese aus der Bedeutung der mentalen Zustände der Erfinder:innen und Nutzer:innen ableitet.

Klarerweise hört die Computer-Analogie hier auf. Wenn wir den Geist erklären wollen, können wir nicht auf Zustände, die rein konventionell zu ihrer Bedeutung kommen, verweisen. Tatsächlich würden wir in einem Erklärungsregress landen: Wer interpretiert die Zustände unseres Gehirns und weist ihnen damit Bedeutung zu? Es scheint, wir müssten einen Homunkulus postulieren (oder eine Seele wie bei Descartes) – also ein kleines Menschlein in unserem Kopf, das diese Interpretationsaufgabe übernimmt. Allerdings hätten wir damit nicht viel erklärt, denn: Wer interpretiert die Zustände des Homunkulus und weist diesen Bedeutungen zu? Die Bedeutung von geistigen Zuständen muss also auf eine andere Weise zustande kommen. Die Philosophie der Kognition verfolgt bis heute ein *Naturalisierungsprogramm* bezüglich des so genannten *repräsentationalen Gehalts* von geistigen Zuständen (d. h. deren Bedeutung). Die Herausforderung lautet: Wie kommt es, dass ein bestimmtes Hirnmuster etwas repräsentiert, unter der Annahme, dass es keinen Homunkulus in meinem Kopf gibt, der dem Hirnmuster Bedeutung, das heißt repräsentationalen Gehalt, zuschreibt? Mit anderen Worten: Findet sich ein plausibler Ansatz, der erklärt, wie neuronale Zustände *nicht-abgeleiteten Gehalt* haben können?

Unter den vielversprechendsten Ansätzen, die eine Antwort auf diese Herausforderung bieten, sind so genannte *teleosemantische Theorien* des mentalen Gehalts (Millikan, 1989; Neander, 1991; Shea, 2018). Laut diesen Theorien bedeutet bzw. repräsentiert ein neuronaler Zustand das, was seine Funktion bestimmt. Anders: Ein bestimmtes Aktivitätsmuster

repräsentiert, sagen wir, Vasen, da das Aktivitätsmuster die Funktion hat, Vasen zu repräsentieren. Auf den ersten Blick scheint sich das Problem hier zunächst nur zu verschieben, denn Funktionen weisen wir physischen Objekten in der Regel zu qua Design: Mein Auto hat die Funktion mich von A nach B zu bringen, da es dafür gebaut wurde. Das Naturalisierungsprogramm des mentalen Gehalts, laut Vertreter:innen von teleosemantischen Theorien, ist also tatsächlich ein Naturalisierungsprogramm des Funktionsbegriffes. Wie kommen physische Entitäten zu ihrer Funktion unabhängig von menschlichen Absichten? Hierauf gibt es verschiedene Antworten, wovon der *evolutionäre teleosemantische Ansatz* die populärste ist: Zumindest für biologische Entitäten gilt, dass deren Funktion durch die biologische Evolution festgelegt wird. Eine biologische Entität X hat die Funktion F in einem Tier T, wenn die Vorfahren von T überlebt haben, da sie X hatten, dass F ausgeführt hat. Zum Beispiel hat die neuronale Aktivität in einem bestimmten Areal eines Froschgehirns die Funktion, Fliegen zu repräsentieren (und damit den repräsentationalen Gehalt/die Bedeutung „Fliege“), wenn die Vorfahren des Frosches überlebt haben, da sie Hirnareale hatten, die Fliegen repräsentiert haben, was die befähigt hat, Nahrung zu finden und zu fangen.

Das Naturalisierungsprogramm des repräsentationalen Gehaltes ist nach wie vor im vollen Gange. Freund:innen des Computationalismus bzw. der Computer-Analogie des Geistes sind optimistisch, dass es erfolgreich sein wird. Der Erfolg dieses Programms hätte den Vorteil, dass wir eine eingängige allgemeine Theorie des Geistes zur Verfügung hätten: Geist besteht in der regelbasierten Manipulation von neuronalen Zuständen, die nicht-abgeleitete Bedeutung tragen. Allerdings sind nicht alle optimistisch, was das Naturalisierungsprogramm angeht. Dies werde ich im Folgenden erläutern.

3 Die Ablehnung der klassischen Kognitionswissenschaft durch die Theorie des situierten Geistes

Die Theorie des situierten Geistes ist weniger eine einheitliche Theorie als ein Forschungsprogramm. In diesem Forschungsprogramm finden sich eine Vielzahl verschiedener Thesen, die nicht immer von allen Autor:innen gleichzeitig und auf gleiche Weise vertreten werden. Ich möchte hier nicht auf spezifische Ansätze eingehen, sondern die Kernmotive des Forschungsprogramms kurz erläutern. Im Sinne der Verständlichkeit werde ich diese Kernmotive so behandeln, als würden sie von den Vertreter:innen dieses Forschungsprogramms einheitlich vertreten. Die Lesenden sollten im Hinterkopf behalten, dass meine Darstellung und Kritik nicht alle Ansätze der situierten Kognition in gleichem Maße betrifft.

Ein Kernmotiv von Vertreter:innen von situierten Ansätzen der Kognition ist, dass sie das Naturalisierungsprogramm des repräsentationalen Gehaltes für gescheitert erachten (Hutto & Myin, 2013). Ein zentrales Problem sehen sie darin, dass es bisher nicht gelungen ist, zu erklären, wie physische Zustände zu einem *spezifischen* Gehalt kommen. Das Problem lässt sich mithilfe des oben erwähnten evolutionären teleosemantischen Ansatzes veranschaulichen: Konnten die Vorfahren des Frosches überleben, weil sie Fliegen repräsentiert haben oder weil sie Nahrung repräsentiert haben oder kleine, schwarze Punkte? Diese Frage lässt sich auf der Basis der evolutionären Teleosemantik nicht beantworten, da die Zuschreibung jeder dieser Gehalte das Überleben des Frosches und seiner Vorfahren erklären würde. An diesem *schwierigen Problem des Gehalts* („hard problem of content“) (Hutto & Myin, 2013) scheitern alle bisherigen Antworten des Naturalisierungsprogramms. Da damit die Versuche von Generationen von Philosoph:innen gescheitert sind, bleibt – so die Vertreter:innen der situierten Ansätze – nicht viel Hoffnung, dass das Problem lösbar ist.

Aus dem Scheitern des Naturalisierungsprogramms des repräsentationalen Gehaltes folgt laut den Vertreter:innen der situierten Kognition auch, dass der Computationalismus nicht haltbar ist. Der Geist und das Gehirn funktionieren nicht wie Software und Hardware. Als Ersatz wird häufig eine andere Analogie bemüht: Der Geist und das Gehirn funktionieren wie ein Fliehkraftregler (Van Gelder, 1995). Der Fliehkraftregler war eine geniale Erfindung, die James Watt im Jahre 1788 einführte, um die Geschwindigkeit der Dampfmaschinen, die zum Beispiel in der Textilproduktion eingesetzt wurden, konstant zu halten. Der Fliehkraftregler kommt ganz ohne Computation aus. Es ist nicht der Fall, dass der Fliehkraftregler einen bedeutungstragenden Input erhält (z. B. die tatsächliche und die gewünschte Geschwindigkeit der Dampfmaschine), um dann auf der Basis eines Algorithmus zu berechnen, ob die Geschwindigkeit der Dampfmaschine gedrosselt oder erhöht werden muss. Die Lösung basiert vielmehr auf einer dynamischen Kopplung der Einzelteile des Fliehkraftregler-Dampfmaschine-Systems: Der Fliehkraftregler ist ein rotierendes Element, das von der Dampfmaschine angetrieben wird. Er besitzt zwei bewegliche Arme, an deren Enden sich jeweils Massen befinden, die sich aufgrund der Fliehkraft, die durch die Rotation zustande kommt, mehr oder weniger stark heben. Die Arme sind zudem mit einer Klappe verbunden, die die Menge des Dampfes regelt, der die Arbeitsgeschwindigkeit bestimmt. Damit bestimmt die Höhe der Auslenkung der Arme, wie stark die Klappe geöffnet oder geschlossen ist. Durch die Kopplung der Hebearme, der Klappe und der Dampfmaschine wird eine konstante Geschwindigkeit erreicht: Sollte zu viel Dampf in das System gelangen, würde der

Fliehkraftregler schneller rotieren, die Arme würden sich heben und die Klappe würde geschlossen – so lange, bis die Arme wieder gesenkt wären.

Wie das Fliehkraftregler-Klappe-Dampfmaschine-System bilden Gehirn, Körper und Umwelt, laut Vertreter:innen der situierten Kognition, ein dynamisches System. Verhalten und Kognition entsteht durch die direkte, dynamische Kopplung zwischen Gehirn, Körper und Umwelt und nicht mittels Repräsentationen.

Hier ein Beispiel aus der kognitiven Psychologie, das diese Kernthese des *Dynamizismus* veranschaulicht: Bittet man Versuchspersonen ein Muster von Legosteinen nachzubauen, so zeigt sich, dass sie dabei nicht die Strategie anwenden, sich das Muster zu merken, um die Kopie zu erstellen. Vielmehr beobachtet man eine dynamische Kopplung zwischen Augen- und Armbewegungen, die so abgestimmt sind, dass die Versuchspersonen die Kopie ohne Fehler erstellen können. Ob sich diese These auf alle kognitiven Aufgaben übertragen lässt, ist Gegenstand einer laufenden Debatte.

Neben der Ablehnung des Repräsentationsbegriffs und des Computationalismus ist ein weiteres Kernmotiv des situierten Forschungsprogramms die Ablehnung des *Internalismus*. Geist findet sich nicht nur im Gehirn, sondern involviert maßgeblich den Körper oder sogar die Umwelt. Diese *These des erweiterten Geistes* ist einerseits motiviert durch ähnliche Überlegungen, wie die zum Fliehkraftregler (Chemero, 2000). Das Gehirn alleine ist nicht der Träger von Kognition – so wie der Fliehkraftregler alleine noch nichts regelt. Kognition involviert das gesamte Gehirn-Körper-Umwelt-System. Durch die dynamische Kopplung der Teile lässt sich eine trennscharfe Unterteilung ohnehin nicht vornehmen. Ein weiteres Argument für den erweiterten Geist ist das sogenannte *Paritätsprinzip* (Clark & Chalmers, 1998): Externe Elemente führen häufig dieselbe Funktion aus wie interne, das heißt neuronale Elemente, die wir als „kognitiv“ bezeichnen. Zum Beispiel kann ein Notizbuch das neuronale Gedächtnis ersetzen, sollte es zum Beispiel durch die Alzheimerkrankheit zerstört sein.

Die bisherige Darstellung zeigt, dass das Forschungsprogramm des situierten Geistes nicht durch eine positive These zur Natur des Geistes zusammengehalten wird. Der gemeinsame Nenner besteht vielmehr in der Kritik an der klassischen Kognitionswissenschaft.

4 Unbewusster Geist?

Nun zurück zu unserem Ausgangsproblem: Was ist das Geistige am „neuen“ Unbewussten? Ein unbestrittenes Kriterium für den Geist ist das Bewusstsein: Jeder bewusste Prozess ist ein geistiger Prozess; jeder bewusste Zustand ein geistiger Zustand. Das Merkmal des Bewusstseins ist also *hinreichend* für die Identifizierung des Geistigen. Die Annahme, dass es auch unbewusste geistige Phänomene gibt, zeigt jedoch, dass Bewusstsein nicht *notwendig* für

das Geistige sein kann. Unsere Herausforderung lautet: Wie lässt sich der unbewusste Geist vom bloß Nicht-Bewussten unterscheiden? Oder in anderen Worten: Gibt es neben dem Bewusstsein ein weiteres hinreichendes Kriterium für den Geist?

Die klassische Kognitionswissenschaft scheint eine einfache Antwort zu haben: Das Geistige besteht in der regelbasierten Manipulation von physischen Zuständen, die nicht-abgeleiteten repräsentationalen Gehalt haben. Dieses Kriterium ist klarer Weise unabhängig davon, ob ein Prozess bewusst abläuft oder nicht. Der unbewusste Geist besteht, wie der bewusste Geist, aus computationalen Prozessen – nur eben ohne Bewusstsein.

Den situierten Ansätzen steht diese Antwort offensichtlich nicht zur Verfügung, wenn sie die Existenz von Repräsentationen bestreiten.⁷ Wie also können diese den Unterschied zwischen dem unbewussten Geist und dem bloß Nicht-Bewussten beschreiben?

Der kurze Überblick über die Kernmotive des Forschungsprogramms der situierten Kognition im vorangegangenen Abschnitt hat verdeutlicht: Im Kern des Programms steht keine klare Definition dessen, was Kognition bzw. den Geist ausmacht. Allerdings fehlt nicht nur eine eindeutige, verbindende Definition des Geistes – es finden sich auch keine lokalen Thesen bezüglich der Natur des Geistes, die für die hier relevante Herausforderung der Abgrenzung des unbewussten Geistes vom bloß Nicht-Bewussten hilfreich wäre. Die Charakterisierung von kognitiven Prozessen als dynamische Systeme von gekoppelten Sub-Systemen bringt uns nicht weiter, da auch Systeme, wie zum Beispiel Turbolader und Motor ein gekoppeltes dynamisches System bilden, das klarerweise kein Teil des Geistes ist. Dynamische Kopplung ist als hinreichendes Kriterium für den Geist also nutzlos.

Eventuell bringt uns das Paritätsprinzip weiter. Das mit dem Prinzip verbundene Kriterium lautet: Würde ein Prozess im Kopf stattfinden und würden wir ihn dann als „geistig“ bezeichnen, dann sollten wir den Prozess auch als „geistig“ bezeichnen, wenn er nicht im Kopf stattfindet (Clark & Chalmers, 1998). Man könnte das Paritätsprinzip entsprechend anpassen: Würde der Prozess bewusst ablaufen und würden wir ihn dann als „geistig“ bezeichnen, dann sollten wir den Prozess auch als „geistig“ bezeichnen, wenn er nicht bewusst abläuft. Das Problem, dass sich dabei ergibt, ist, dass Bewusstsein bereits als hinreichendes Kriterium für das Geistige gilt (im Gegensatz zu „im Kopf stattfinden“) und damit stellt sich das so umformulierte Kriterium als trivial heraus: Natürlich würden wir einen Prozess, der bewusst abläuft, als „geistig“ bezeichnen – eben, weil er bewusst ist. Daraus folgt nicht, dass wir den Prozess auch dann als „geistig“ bezeichnen würden, wenn er nicht bewusst ablaufen würde

⁷ Einige Vertreter:innen des situierten Ansatzes akzeptieren die Existenz von mentalen Repräsentationen (z.B. Rowlands (2009)).

bzw. ist das ja genau die Frage, um die es hier geht. Eine alternative Interpretation des Kriteriums könnte dieses Problem vermeiden: wenn wir es als eine Aussage über die funktionale Ähnlichkeit zwischen klarerweise geistigen Prozessen und jenen, deren Status als geistig geklärt werden soll, verstehen. Anders ausgedrückt: Wenn ein unbewusster Prozess die gleiche funktionale Rolle hat wie ein bewusster Prozess, dann sollte der unbewusste Prozess auch als „geistig“ bezeichnet werden.

Allerdings entsteht hier ein Problem, das dem ähnelt, das auch im Kontext der These des erweiterten Geistes diskutiert wird (Rupert, 2004): Individuiert man die funktionalen Merkmale der klarerweise geistigen Prozesse grobkörnig, dann sind die funktionalen Rollen der relevanten unbewussten Prozess wahrscheinlich identisch. Zum Beispiel: Wenn wir bewusstes Sehen charakterisieren als die Verarbeitung von visuellen Reizen, die zu Verhaltensänderungen führt, dann haben Fälle von unbewusstem Sehen klarerweise dieselben funktionale Rolle. Allerdings wird diese grobkörnige Beschreibung der tatsächlichen funktionalen Rolle des bewussten Sehens nicht gerecht, da bewusstes Sehen sich nicht in dieser funktionalen Beschreibung erschöpft (z. B. führt bewusstes Sehen auch dazu, dass wir über das Gesehene berichten können). Wenn man aber eine feinkörnigere Beschreibung wählt, die der tatsächlichen funktionalen Rolle des bewussten Sehens gerecht wird, so hätte unbewusstes Sehen klarerweise eine andere funktionale Rolle und würde somit nicht als „Sehen“, also als geistig, zählen.

Es bleibt also offen, ob sich die Redeweise vom unbewussten Geist im Kontext situierter Theorien des Geistes beschreiben lässt. Im Folgenden möchte ich eine Idee skizzieren, auf deren Basis eine Unterscheidung zwischen dem unbewussten Geist und dem bloß Nicht-Bewussten auch ohne den Verweis auf mentale Repräsentationen gemacht werden könnte.

5 Ein wissenschaftstheoretischer Blick auf die Herausforderung

Die zentrale Beobachtung, die meiner Idee zugrunde liegt, ist die folgende: Der unbewusste Geist ist ein Postulat der gegenwärtigen empirischen Psychologie, das dazu dient, bestimmte in Experimenten beobachtete Verhaltensweisen zu erklären. Nehmen wir das Beispiel des unbewussten Sehens im Kontext der Binokularen Rivalität wie in der Einleitung beschrieben. Das Verhalten, das Versuchspersonen zeigen, lässt sich wie folgt charakterisieren: Erstens äußern sie, dass sie die Bilder der nackten Personen nicht sehen. Zweitens reagieren sie auf Reize, die an der Stelle der Bilder nach deren Präsentation auftauchen, schneller als auf Bilder, die an einer anderen Stelle auftauchen (je nach sexueller Präferenz), woraus geschlossen wird, dass die Aufmerksamkeit der Versuchspersonen von den Bildern der Nackten beeinflusst wird. Die Frage, die Psycholog:innen nun beantworten wollen, lautet: Wie kann der zweite Punkt

erklärt werden, gegeben dem ersten? Die Antwort der meisten Psycholog:innen lautet: Die Erklärung ist, dass die Versuchspersonen die Bilder der Nackten unbewusst sehen.

Der unbewusste Geist soll also Verhalten erklären. Die Existenz des unbewussten Geistes wird aus einem *Schluss auf die beste Erklärung* hergeleitet. Die Aufmerksamkeitsbeeinflussung in den Experimenten kann am *besten* dadurch erklärt werden, dass die Versuchspersonen die Bilder der Nackten unbewusst sehen. Ob dieser Schluss gerechtfertigt ist, hängt davon ab, ob der unbewusste Geist tatsächlich die beste Erklärung für das zu erklärende Verhalten liefert. Meine These ist, dass die Beantwortung dieser Frage der Analyse des Begriffs des unbewussten Geistes zugrunde liegen sollte. In anderen Worten: Wie muss der Begriff des unbewussten Geistes verstanden werden, damit das Postulat eines unbewussten mentalen Phänomens potentiell als beste Erklärung gelten kann?

Um diese Frage zu beantworten, können wir uns den Ressourcen der gegenwärtigen Wissenschaftstheorie bedienen. Eine prominente These lautet, dass wissenschaftliche Erklärungen in der Psychologie *mechanistische Erklärungen* sind (Wright & Bechtel, 2007). Mechanistische Erklärungen zeigen, wie das Verhalten eines Systems durch die Organisation und Interaktion der Systemteile, also durch den zugrundeliegenden Mechanismus, zustande kommt. Erklärungen, die auf den unbewussten Geist verweisen, sind also als mechanistische Erklärungen zu verstehen. Diese Erklärungen beschreiben Mechanismen, die unbewusste mentale Prozesse als Komponenten haben. Die oben gestellte Frage lässt sich also umformulieren: Wie muss der Begriff des unbewussten Geistes verstanden werden, damit unbewusste geistige Prozesse potentiell Komponenten von Mechanismen sein können?

Eine weitere These zur wissenschaftlichen Erklärung lautet, dass Erklärungen kontrastive Fragen beantworten (Craver & Kaplan, 2020; Kohár & Krickel, 2021). Eine Erklärung des unbewussten Sehens zum Beispiel muss die Frage beantworten, warum die nicht bewusst wahrgenommenen Stimuli die Aufmerksamkeit beeinflussen und nicht etwa keine Auswirkung haben. Meine These ist: Wenn die mechanistische Erklärung, die diese Frage beantwortet, einen Mechanismus beschreibt, der auch die Aufmerksamkeitsbeeinflussung bei bewusstem Sehen erklären würde, dann haben wir es mit einem Fall von unbewusstem *Sehen*, also einem geistigen Prozess zu tun. Wir müssen also den Mechanismus des bewussten Sehens mit dem Mechanismus des mutmaßlich unbewussten Sehens vergleichen und herausfinden, ob sich diese in relevanter Weise ähneln, das heißt bezüglich jener Komponenten, die für die Aufmerksamkeitsbeeinflussung verantwortlich sind. Ob dies der Fall ist, muss auf Basis der empirischen Evidenz geklärt werden.

Meine These lautet also: Der unbewusste Geist unterscheidet sich vom bloß Nicht-Bewussten darin, dass die Mechanismen, die unbewussten geistigen Prozessen zugrunde liegen, den Mechanismen ähneln, die bewussten geistigen Prozessen zugrunde liegen. Dieses Kriterium ist mit allen wissenschaftlichen Theorien des Geistes kompatibel – auch jenen, die die Existenz von mentalen Repräsentationen bestreiten, also auch Theorien des situierten Geistes.

6 Schluss

Die Herausforderung für Theorien des Geistes im Kontext des unbewussten Geistes lautet: Was unterscheidet den unbewussten Geist vom bloß Nicht-Bewussten? Klassische Theorien des Geistes haben eine klare Antwort: Geistige Prozesse bestehen in der regelbasierten Manipulation von bedeutungstragenden physikalischen Zuständen. Diese können bewusst oder unbewusst ablaufen. Neuere Ansätze der situierten Kognition können die Herausforderung nicht so einfach beantworten aus drei Gründen: Erstens wird die Existenz von mentalen Repräsentationen (bzw. bedeutungstragenden physischen Zuständen) von (den meisten) Vertreter:innen der situierten Kognition bestritten. Zweitens haben die unterschiedlichen Ansätze zur situierten Kognition unterschiedliche Ansichten darüber, was den Geist auszeichnet. Drittens ist keines der bisher vorgeschlagenen nicht-repräsentationalen Kriterien für das Geistige (dynamische Kopplung, funktionale Parität) ausreichend, um eine Unterscheidung zwischen unbewusstem Geist und bloß Nicht-Bewussten zu treffen.

In diesem Aufsatz habe ich eine Lösungsstrategie vorgeschlagen, die helfen könnte, alle diese Hürden zu überwinden. Die Lösung ist unabhängig davon, welche konkrete Theorie des Geistes vertreten wird und kann somit zum Beispiel auch von Vertreter:innen des Dynamizismus und Vertreter:innen der These des erweiterten Geistes akzeptiert werden. Ein weiterer Vorteil der vorgeschlagenen Lösungsstrategie ist, dass sie die Wissenschaft des Geistes als Ausgangspunkt nimmt. Der unbewusste Geist wird als ein empirisch-explanatorisches Postulat begriffen, dessen Existenz durch einen Schluss auf die beste Erklärung von empirischen Daten gerechtfertigt ist. Eine offene Frage, die es zu klären gilt, ist ob und wie sich diese Lösungsstrategie auf den Alltagsbegriff des unbewussten Geistes übertragen lässt. Wenn wir im Alltag das Verhalten unserer Mitmenschen erklären, indem wir ihnen unbewusste mentale Zustände zuschreiben, haben wir dabei in der Regel keine mechanistischen Erklärungen im Sinn. Vielmehr geht es uns um rechtfertigende Erklärungen, die das Verhalten einer Person nachvollziehbar machen sollen. Die Unterschiede in der Erklärungspraxis zwischen Wissenschaft und Alltag legen nahe, dass es zwei Begriffe des unbewussten Geistes gibt – einen wissenschaftlichen und einen alltäglichen. Zu untersuchen

wäre, wie genau sich die Begriffe unterscheiden, ob die Unterschiedlichkeit zu Missverständnissen führt und ob Wissenschaftler:innen diese Begriffe verwechseln könnten, wodurch es zu Fehlinterpretationen von Daten kommen könnte.

Referenzen

- Ayan, S. (2018). Das Unbewusste - Warum es noch viel mächtiger ist als wir ahnen. *Gehirn & Geist*, 10.
- Banaji, M. R., Bhaskar, R. & Brownstein, M. (2015). When bias is implicit, how might we think about repairing harm? *Current Opinion in Psychology*, 6, 183–188. Elsevier Ltd.
- Bargh, J. (2017). *Before You Know It: The Unconscious Reasons We Do What We Do*. Random House. Verfügbar unter: <https://books.google.de/books?id=GSitDQAAQBAJ>
- Berlin, H. A. (2011). The neural basis of the dynamic unconscious. *Neuropsychanalysis*, 13(1), 5–31. <https://doi.org/10.1080/15294145.2011.10773654>
- Bushman, B. J., Dwall, C. N., Pond, R. S. & Hanus, M. D. (2014). Low glucose relates to greater aggression in married couples. *Proceedings of the National Academy of Sciences of the United States of America*, 111(17), 6254–7. National Academy of Sciences. <https://doi.org/10.1073/pnas.1400619111>
- Chemero, A. (2000). Anti-Representationalism and the Dynamical Stance. *Philosophy of Science*, 67(4), 625–647.
- Clark, A. & Chalmers, D. J. (1998). The extended mind. *Analysis*, 58(1), 7–19.
- Craver, C. F. & Kaplan, D. M. (2020). Are More Details Better? On the Norms of Completeness for Mechanistic Explanations. *The British Journal for the Philosophy of Science*, 71, 287–319. <https://doi.org/10.1093/bjps/axy015>
- Van Gelder, T. (1995). What Might Cognition Be, If Not Computation? *The Journal of Philosophy*, 92(7), 345–381.
- Greenwald, A. G., Mcghee, D. E. & Schwartz, J. L. K. (1998). Measuring Individual Differences in Implicit Cognition: The Implicit Association Test. *Journal of Personality and Social Psychology*, 74(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>
- Hahn, A., Judd, C. M., Hirsh, H. K. & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*, 143(3), 1369–1392. <https://doi.org/10.1037/a0035028>
- Hassin, R. R. (2013). Yes It Can: On the Functional Abilities of the Human Unconscious. *Perspectives on Psychological Science*, 8(2), 195–207. SAGE PublicationsSage CA: Los Angeles, CA. <https://doi.org/10.1177/1745691612460684>
- Hassin, R. R., Uleman, J. S. & Bargh, J. A. (2005). *The New Unconscious. The New*

- Unconscious*. New York: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780195307696.001.0001>
- Hutto, D. D. & Myin, E. (2013). *Radicalizing Enactivism: Basic Minds Without Content*. Cambridge: MIT Press.
- Jiang, Y., Costello, P., Fang, F., Huang, M. & He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences*, 103(45), 17048–17052.
<https://doi.org/10.1073/pnas.0605678103>
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Farrar, Straus and Giroux.
- Kohár, M. & Krickel, B. (2021). Compare and Contrast: How to Assess the Completeness of Mechanistic Explanation. In F. Calzavarini & M. Viola (Hrsg.), *Neural Mechanisms - New Challenges in the Philosophy of Neuroscience* (S. 395–424). Springer.
https://doi.org/10.1007/978-3-030-54092-0_17
- Krickel, B. (2018). Are the states underlying implicit biases unconscious? – A Neo-Freudian answer. *Philosophical Psychology*, 31(7), 1007–1026.
<https://doi.org/10.1080/09515089.2018.1470323>
- Millikan, R. G. (1989). In Defense of Proper Functions. *Philosophy of Science*, 56(June), 288–302. University of Chicago Press.
- Mlodinow, L. (2013). *Subliminal: how your unconscious mind rules your behavior*. Vintage Books.
- Neander, K. (1991). The teleological notion of ‘function’. *Australasian Journal of Philosophy*, 69(4), 454–468. <https://doi.org/10.1080/00048409112344881>
- Pernu, T. K. (2017). The five marks of the mental. *Frontiers in Psychology*, 8(JUL), 1084. Frontiers. <https://doi.org/10.3389/fpsyg.2017.01084>
- Rowlands, M. (2009). Extended cognition and the mark of the cognitive. *Philosophical Psychology*, 22(1), 1–19. <https://doi.org/10.1080/09515080802703620>
- Rupert, R. D. (2004). Challenges to the hypothesis of extended cognition. *Journal of Philosophy*, 101(8), 389–428. <https://doi.org/10.5840/jphil2004101826>
- Shea, N. (2018). Representation in Cognitive Science.
- Weiskrantz, L. (1995). Blindsight: Not an Island Unto Itself. *Current Directions in Psychological Science*, 4(5), 146–151.
- Wright, C. & Bechtel, W. (2007). Mechanisms and Psychological Explanation. *Philosophy of Psychology and Cognitive Science* (S. 31–79). Elsevier. <https://doi.org/10.1016/B978-044451540-7/50019-0>