

Der nachfolgende Text ist im Jahr 2008 unter dem Titel *Funktionen des Bewusstseins* in der Reihe „Humanprojekt – Zur Stellung des Menschen in der Natur“ im Verlag Walter de Gruyter GmbH & Co KG erschienen.

Bewusstsein als Korrektiv

ALEXANDRA M. FREUND/KLAUS OBERAUER

Die zentrale These dieses Beitrages ist, dass eine Funktion von Bewusstsein darin besteht, als Korrektiv von automatischen Kognitionen und Verhaltensweisen zu wirken, wenn diese gestört werden oder nicht angemessen sind. Meist werden uns diejenigen internen oder externen Aspekte der Welt bewusst, die automatische Prozesse stören, verhindern oder dysfunktional machen. Bewusstsein setzt also ein, wenn es einen Konflikt gibt zwischen den automatisch ablaufenden Verhaltensabfolgen und den situativen Gegebenheiten oder den jeweiligen Zielen einer Person. Mit Korrektiv ist gemeint, dass die mit dem Bewusstsein einhergehende Kontrollierbarkeit von Verhalten es Personen ermöglicht, ihr Verhalten an die situativen Gegebenheiten und ihre Ziele anzupassen. Die Alternativ-These besteht darin, dass Bewusstsein lediglich ein Epiphänomen von korrektiven Kontrollprozessen darstellt, selbst jedoch keinerlei Funktion besitzt. Um zwischen diesen beiden Modellen auf einer empirisch-psychologischen Ebene entscheiden zu können, wird in der Psychologie experimentell getestet, ob sich Verhalten in Abhängigkeit davon unterscheidet, ob sich eine Person eines bestimmten Stimulus und/oder Verhaltens bewusst ist. Hierzu werden beispielhaft ausgewählte empirische Studien referiert, die als indirekter Beleg für die These von Bewusstsein als Korrektiv angesehen werden können.

Erforschung von Bewusstsein in der Psychologie

Wie kann in der Psychologie Bewusstsein erforscht werden? Eine nahe liegende Herangehensweise ist, Personen um eine Mitteilung darüber zu bitten, ob sie etwas bewusst erleben. Diese Herangehensweise ist für die Untersuchung der Inhalte des Geistes (z.[^]B. Wahrnehmungen, Gedanken, Gefühle) durchaus üblich und sinnvoll. Wir können eine Person fragen, welches Wort sie auf einem Computerbildschirm gesehen hat, und wenn sie das Wort richtig angeben kann, können wir schließen, dass sie das Wort tatsächlich wahrgenommen hat – anders wäre ihr Verhalten kaum zu erklären. Was eine Person uns mitteilt, gibt uns also einen empirischen Zugang zu den Inhalten ihrer Kognition, aber können wir in gleicher Weise auch Auskunft darüber bekommen, ob eine Wahrnehmung, ein Gedanke, ein Gefühl bewusst ist? Der Begriff des Bewusstseins impliziert eine bestimmte subjektive *Erlebensqualität*, die mit einer Wahrnehmung, einem Gefühl oder einem

Gedanken einhergeht. Diese Erlebensqualität ist von außen prinzipiell nicht beobachtbar. Wenn eine Person sagt, sie habe das Wort „Philosophie“ bewusst gesehen, können wir daraus schließen, dass sie eine Repräsentation des Wortes hatte (und auch, dass sie eine Repräsentation der Aussage: „Ich habe dieses Wort bewusst gesehen“ hatte), aber nicht, dass diese Repräsentation tatsächlich von der subjektiven Qualität des Bewusstseins begleitet war. Dieselbe Repräsentation ohne Bewusstsein hätte dasselbe Verhalten, dieselbe Mitteilung hervorgebracht, und daher eröffnet das beobachtbare Verhalten keinen empirischen Zugang zu der subjektiven Qualität mentaler Zustände.

In der Philosophie des Bewusstseins wird häufig unterschieden zwischen phänomenalem Bewusstsein einerseits und Zugriffsbewusstsein (Block 1995) oder psychologischem Bewusstsein (Chalmers 1996) andererseits. Phänomenales Bewusstsein bezieht sich auf die subjektive Erlebensqualität mentaler Zustände (z.[^]B. Wahrnehmungen, Gefühle oder Gedanken). Zugriffs- oder psychologisches Bewusstsein ist definiert durch die kausale Rolle, die ein bewusster Zustand für mentale Prozesse und Verhalten hat. Das oben angesprochene Problem der fehlenden empirischen Indikatoren für Bewusstsein betrifft das phänomenale Bewusstsein. Die empirische Psychologie des Bewusstseins ignoriert dieses Problem weitgehend. In der Forschungspraxis wird Bewusstsein in der Regel durch Mittelbarkeit operationalisiert. Ein Gedanke oder Gefühl gilt dann als bewusst, wenn eine Person darüber Auskunft geben kann. Dies kann verbal oder non-verbal sein (z.[^]B. Knopfdruck beim Erkennen eines bestimmten Stimulus). Eine Mitteilung kann von außen beobachtet werden und ist damit der intersubjektiven Überprüfbarkeit zugänglich. Diese Operationalisierung bezieht sich (oft implizit) auf das Zugriffs- oder psychologische Bewusstsein, weil ein mentaler Zustand dann und nur dann mittelbar ist, wenn er kausal Einfluss auf das Verhalten hat, das wir als Mitteilung interpretieren. Die Psychologie beschränkt sich notwendigerweise auf die Erforschung dieser Art des Bewusstseins, da das phänomenale Bewusstsein grundsätzlich einer intersubjektiven, empirischen Untersuchung nicht zugänglich ist.

Ein Problem des Mittelbarkeitskriteriums besteht darin, dass die Mitteilung von Bewusstseinsinhalten der subjektiven Einschätzung von „Bewusstheit“ unterliegt. Man weiß nicht, welche Kriterien eine Person anlegt, um Bewusstheit bei sich festzustellen. Subjektive Berichte von Bewusstheit reflektieren daher neben dem mentalen Zustand, über den berichtet wird, auch die naiven Theorien von Bewusstheit und welche Phänomene hierfür indikativ seien. Zum Beispiel kann man ein Wort nur wenige Millisekunden darbieten (gefolgt von einer visuellen Maske, die das so genannte Nachbild löscht) und die Person dann fragen, ob sie das Wort bewusst wahrgenommen hat. Manche würden vielleicht den vagen Eindruck von Buchstabenfragmenten als bewusste Wahrnehmung des Wortes auffassen, während für andere dies nicht ausreichend wäre. Als eine Lösung dieses Problems wurde vorgeschlagen, Personen zwischen Situationen diskriminieren zu lassen, in denen ein Stimulus dargeboten wurde, und solchen, in denen kein Stimulus gezeigt wurde (z.[^]B. nur die visuelle Maske ohne vorangehendes Wort). Verhaltenskriterium für Bewusstsein ist dann eine Diskriminationsleistung, die über der Ratewahrscheinlichkeit liegt. Allerdings hat man hier letztendlich dasselbe Problem, nämlich dass Personen unterschiedliche Kriterien anlegen, um zu bestimmen, ob sie etwas wahrgenommen haben oder nicht. Daher ist der Königsweg nach Merikle und Daneman (1998) zu zeigen, dass Stimuli, von denen eine Person in Wort oder Tat angibt, sie nicht bewusst wahrgenommen zu haben, eine andere psychologische Wirkung haben als solche, die bewusst wahrgenommen wurden. Dies wird in der Psychologie als Dissoziation bezeichnet. Hierauf gehen wir später noch genauer ein.

Ein weiteres Problem bei dem Kriterium der Mittelbarkeit ist, dass es, gemessen an unserem intuitiven Verständnis von Bewusstsein, zu inklusiv ist. Bei einem als Mitteilung interpretierten Verhalten kann es sich um ein rein reflexhaftes oder programmiertes Verhalten handeln. So würde man beispielsweise einem Lichtsensor, der so programmiert ist, dass beim Lichteinfall ein Sprachsynthesizer mit dem Wort „Licht“ reagiert, kein Bewusstsein zuschreiben (vgl. Block 1981, für eine ausführliche Diskussion in Bezug auf das Bestehen des Turing-Testes und die Zuschreibung von Intelligenz). Ein auf Menschen bezogenes Beispiel ist, dass eine Person, die per klassischer

Konditionierung darauf trainiert wurde, auf einen Glockenschlag mit Lidschlag zu reagieren, den Glockenschlag nicht bewusst wahrzunehmen braucht, um mit einem Lidschlag zu reagieren. Zwar handelt es sich dabei ganz eindeutig um eine Reaktion auf den Glockenschlag, die man als Mitteilung über die Wahrnehmung des Glockentons interpretieren kann. Dennoch würde dies kaum jemanden davon überzeugen, dass die Person den Glockenton bewusst wahrnimmt. Wenn die Person jedoch wahlweise auf den Ton auch mit Handheben oder der Äußerung des Wortes „Glocke“ reagieren kann, so wäre dies ein stärkeres Argument für die Zuschreibung vom Bewusstsein des Glockentons. Wir schlagen daher als in der Psychologie gebräuchliche Arbeitsdefinition von Bewusstsein vor, dass ein Organismus Freiheitsgrade in seiner Reaktion auf bestimmte (innere oder äußere) Reize hat, die er beispielsweise dazu nutzen kann, einen Reiz in verschiedener Weise mitzuteilen.

Was heißt es, Freiheitsgrade in der Reaktion auf Reize zu haben? In Übereinstimmung mit Bieri (2001) fassen wir Freiheitsgrade nicht als das Freisein von kausaler Determination auf, sondern so, dass die Reaktion auf einen Reiz durch außerhalb der Repräsentation des Reizes liegende, andere Repräsentationen moderiert wird (z.[^]B. das Ziel, die Wahrnehmung des Stimulus „Glockenschlag“ mit dem Lidschlag, dem Heben einer Hand oder der Äußerung des Wortes „Glocke“ mitzuteilen). Wir meinen also nicht Freiheitsgrade im Sinne der Unbestimmtheit, sondern der multiplen Bestimmtheit, die (mehr oder weniger) flexible Antworten erlaubt. Eine bewusste Repräsentation ist dieser Auffassung zufolge also eine, bei der Deliberation möglich ist. Deliberation bedeutet, dass die Repräsentation des Stimulus mit einer Vielzahl anderer Repräsentationen in Beziehung gesetzt werden kann. Aus einer starren Reiz-Reaktions-Verknüpfung wird damit ein flexibleres Reaktionsrepertoire, das unterschiedliche Reaktionsmöglichkeiten auf denselben Reiz erlaubt.

Diese begriffliche Rekonstruktion der in der Psychologie üblichen Arbeitsdefinition von Bewusstsein führt uns daher zu einer ersten Antwort auf die Frage nach der Funktion des Bewusstseins. Eine Funktion von Bewusstsein ist dieser Perspektive zufolge, ein flexibleres Reaktionsrepertoire zu schaffen. Das Verhalten wird dann nicht notwendigerweise von einem auslösenden Stimulus gesteuert, sondern von einer neuen Verbindung zwischen dem Stimulus und einer oder mehreren anderen Repräsentationen. Solche Repräsentationen können beispielsweise die Erinnerung an frühere oder die Antizipation künftiger Ereignisse oder Ziele sein. Die Einbeziehung von Zielen ist besonders wichtig, da die Flexibilisierung des Verhaltens kein Selbstzweck ist, sondern der Optimierung der Passung zwischen Umweltgegebenheit und den Zielen des Organismus dient.

Ziele werden in der Psychologie definiert als kognitive Repräsentationen von erwünschten (oder unerwünschten) Zuständen, die ein Organismus anstrebt (oder vermeidet; z.[^]B. Kruglanski 1996). Wird ein Stimulus mit Zielen in Beziehung gesetzt, können Personen auf flexible Weise gemäß ihren Zielen handeln, statt mit einem stereotypen Verhalten zu reagieren. Dies bedeutet nicht, dass eine Person immer gemäß ihren Zielen handelt, sondern dass sie dies prinzipiell *kann*. Es bedeutet weiterhin nicht, dass Ziele bewusst (im Sinne des psychologischen oder Zugriffsbewusstseins) sein müssen, um verhaltensrelevant zu sein. Wie wir im Folgenden zeigen werden, können Ziele durchaus nicht-bewusst sein. Wir postulieren jedoch, dass bewusste Ziele zu einem Gewinn an Flexibilität führen, der eine Korrektur von automatischem Verhalten auf einen gegebenen Reiz erlaubt.

Vorschlag eines hierarchischen Handlungsmodells

In diesem Abschnitt schlagen wir ein hierarchisches Handlungsmodell vor, nach dem Verhalten auf drei Weisen reguliert werden kann: (1) durch direkte Reiz-Reaktionsverbindungen (das Verhalten wird reflexartig durch einen Reiz ausgelöst), (2) durch situative Reize, die automatisch Ziele und damit verbundene Verhaltensabläufe aktivieren, und (3) durch bewusste Ziele, die die Auswahl von

zielbezogenen Handlungen leiten. Die Funktion von Bewusstsein lässt sich am besten durch die Gegenüberstellung der beiden letzteren zeigen. In beiden Fällen werden Ziele als vermittelnde Prozesse der Handlungssteuerung berücksichtigt. Der einzige Unterschied besteht darin, ob diese bewusst sind. Unsere These ist, dass der bewusste Regulationsmodus (3) nur dann in Kraft tritt, wenn die automatischen Modi (1 und 2) aufgrund von Störungen oder Konflikten nicht reibungslos ablaufen.

Das hierarchische Handlungsmodell basiert auf den in der Psychologie gegenwärtig sehr verbreiteten, so genannten Zwei-Prozess-Modellen, die in leicht unterschiedlicher Form alle dieselbe Aussage machen, nämlich dass man zwei Arten der Informationsverarbeitung und des Verhaltens voneinander unterscheiden kann: (1) Nicht-bewusste Prozesse und (2) bewusste Prozesse. Diese beiden Arten der Informationsverarbeitung sind in Tabelle 1 hinsichtlich ihrer zentralen Charakteristika zusammenfassend gegenüber gestellt (vgl. Chaiken und Trope 1999).

Tab. 1: Gegenüberstellung von bewussten und nicht-bewussten Informationsverarbeitungsprozessen nach Zwei-Prozess-Modellen (Chaiken und Trope 1999).

Nicht-bewusste Prozesse	Bewusste Prozesse
Assoziativ ("spreading activation")	Anwendung von (propositionalem) Wissen
Nicht intentional	Intentional
Automatisch (nicht kontrolliert)	Kontrolliert
Schnell	Langsam
Unflexibel	Flexibel
Geringe kognitive Kapazitätsanforderungen	Hohe kognitive Kapazitätsanforderungen

Aufbauend auf den Zwei-Prozess-Modellen, gehen wir davon aus, dass Verhalten per Voreinstellung (und falls möglich) über die ressourcensparenden automatischen Prozesse reguliert wird. Diese können entweder in (reflexhaften) Reiz-Reaktionsverbindungen bestehen oder in der automatischen Aktivierung einer Verbindung zwischen Reiz und Ziel, das dann wiederum automatisch das Verhalten steuert.

Der Begriff der automatisierten Ziele wurde von Bargh (1990) eingeführt. Bargh geht davon aus, dass Ziele genauso wie andere Reize qua klassischer Konditionierung – durch wiederholte gemeinsame Aktivierung – an bestimmte auslösende Reize gebunden werden können. Auf diese Weise kann ein bestimmter Reiz ein Ziel automatisch aktivieren, d.h. ohne dass deliberativ eine Verbindung der Repräsentation des Reizes mit einer Zielrepräsentation herbeigeführt werden muss. Das solchermaßen automatisch aktivierte Ziel kann nun seinerseits wiederum automatisch Verhalten aktivieren. Der Prozess ist hierbei derselbe: Durch wiederholte, gleichzeitige Aktivierung eines Zieles (Reiz) und eines nachfolgenden Verhaltens (Reaktion), entsteht eine so enge Kopplung zwischen Ziel und Reaktion, dass sie automatisch wird. Nach diesem Modell können also Ziele automatisch durch Umweltreize aktiviert werden, die dann ihrerseits automatisch bestimmtes Verhalten nach sich ziehen.

Eine Reihe empirischer Untersuchungen stützen dieses Modell. Hier soll nur ein Beispiel kurz umrissen werden. So zeigen Fitzsimmons und Bargh (2003) in einer Serie von Studien, dass die Beantwortung von Fragen, die sich auf einen Freund beziehen, in höherem Maße das Ziel „Hilfsbereitschaft“ aktivieren als die Beantwortung derselben Fragen in Bezug auf einen Arbeitskollegen. In einer zweiten Studie mit anderen Versuchsteilnehmenden zeigte sich, dass Personen, die zuvor Fragen in Bezug auf einen Freund beantworteten, einem Fremden gegenüber signifikant mehr Hilfsbereitschaft zeigten als Personen, die zuvor dieselben Fragen in Bezug auf einen Arbeitskollegen beantwortet hatten. Der Inhalt der Fragen hatte nichts mit dem Thema Hilfsbereitschaft zu tun, und keiner der Teilnehmenden berichtete, dass die Verbindung zwischen dem Ausfüllen des Fragebogens und der Hilfsbereitschaft bewusst war. Die Aktivierung des Konzeptes „Freundschaft“ (durch das Ausfüllen des entsprechenden Fragebogens) schien zu der automatischen Aktivierung des Ziels „Hilfsbereitschaft“ geführt zu haben, das wiederum automatisch hilfsbereites Verhalten aktivierte. **Abbildung 1** verdeutlicht diesen Ablauf.

Abb. 1: Automatische Prozesse der Ziel- und Verhaltensaktivierung.

Solange solche automatischen Aktivierungs- und Verhaltensabläufe eine reibungslose Interaktion mit der Umwelt ermöglichen, sollten diese Prozesse aufgrund ihrer besonderen Effizienz – sie laufen schnell ab und stellen nur sehr geringe Anforderungen an die kognitive Kapazität – gewissermaßen als Voreinstellung dienen. Dieses Prinzip bezeichnen wir als das Primat der automatischen Prozesse. Nur wenn automatische Prozesse zu Problemen der Passung eines Organismus und den Anforderungen der Umwelt führen, sollte der bewusste Verarbeitungsmodus in Kraft treten. Probleme kann es zum einen geben, wenn ein Reiz ambig ist (d.h. nicht eindeutig mit einem bestimmten Ziel verknüpft ist). Dann ist der direkte Pfad vom Stimulus zum Ziel nicht mehr gegeben. Dies führt dazu, dass deliberative Prozesse der Zielauswahl stattfinden.

Abb. 2: Deliberative Prozesse der Zielauswahl aufgrund eines uneindeutigen Stimulus.

Zum anderen kann der automatische Prozess dann gestört sein, wenn ein Ziel nicht eindeutig mit einem Verhalten assoziiert ist (z.B. weil die Kopplung von Ziel und Verhalten noch nicht stark genug ist, um zu einer automatischen Verhaltensaktivierung zu führen). In diesem Fall werden deliberative Prozesse der Handlungsauswahl stattfinden, um eine Passung zwischen Ziel und Verhalten herzustellen.

Abb. 3: Deliberative Prozesse der Handlungsauswahl aufgrund eines uneindeutigen Zieles.

Schließlich kann eine Störung der automatischen Verknüpfung von Ziel und Verhalten auch dann auftreten, wenn die Realisierung eines Verhaltens nicht möglich ist, da hierzu die Fertigkeiten nicht vorhanden sind oder dies die Situation nicht erlaubt. Auch in diesem Fall erfolgt eine deliberative Auswahl einer zielbezogenen Handlung.

Abb. 4: Deliberative Prozesse der Handlungsauswahl aufgrund von geblockten Verhaltensabläufen.

Die Unterbrechung einer automatischen Kette von Stimulus, Ziel und Verhalten wird uns meist als eine Entscheidungsmöglichkeit bewusst, in der wir unser propositionales Wissen anwenden können und damit Flexibilität durch das Setzen eines neuen Ziels oder durch eine Handlungsauswahl erlangen.

Ein möglicher Einwand gegen diese (wie auch jede andere) These zur Funktion von Bewusstsein ist, dass die Frage nach der Funktion von Bewusstsein falsch gestellt ist, da sie davon ausgeht, dass Bewusstsein *zu etwas gut sei* und nicht schlicht ein emergentes Phänomen, das entsteht, wenn sich der Organismus in einem bestimmten Zustand befindet. Wie kann man nun zwischen diesen beiden Hypothesen entscheiden? Eine Möglichkeit besteht darin, dass man testet, ob sich Verhalten in Abhängigkeit davon unterscheidet, ob sich Personen eines bestimmten Stimulus und/oder Verhaltens bewusst sind (Mayr 2004). Nur wenn sich zeigen lässt, dass eine solche Dissoziation zwischen bewussten und nicht-bewussten Prozessen besteht, kann darauf geschlossen werden, dass das Bewusstsein kein reines Epiphänomen darstellt. Im Folgenden werden hierzu beispielhaft einige sozialpsychologische Studien referiert, die die These stützen, dass Bewusstsein als Korrektiv für unangemessene oder gestörte Verhaltensabläufe fungiert.

Zunächst soll jedoch kurz geklärt werden, was mit Korrektiv gemeint ist. Wie von einer Reihe von Autoren (z.B. Posner und DiGirolamo 1998) gefunden wurde, werden uns meist diejenigen internen oder externen Aspekte der Welt bewusst, die Routine-Prozesse stören, verhindern oder dysfunktional machen. Bewusstsein tritt also dann auf den Plan, wenn es einen Konflikt zwischen automatisch ablaufenden Handlungsabfolgen und den situativen Anforderungen oder den Zielen der Person gibt. Mit Korrektiv ist also gemeint, dass die mit dem Bewusstsein einhergehende Kontrollierbarkeit von Verhalten dazu führt, dass Personen ihre Ziele und Verhalten im Sinne ihrer übergeordneten Ziele an die situativen Gegebenheiten anpassen können. Ob dies immer zu den für die Person besten Ergebnissen führt, ist für dieses Argument nebensächlich (wie Dijksterhuis und Nordgren (2006) in einem Übersichtsartikel auf der Grundlage empirischer Evidenz argumentieren, können diese Korrekturversuche aufgrund unserer limitierten kognitiven Kapazität durchaus zu schlechteren Ergebnissen führen als nicht-bewusste Vorgänge).

Empirische Beispiele

Im Folgenden soll beispielhaft über einige empirische Studien berichtet werden, die die These vom Bewusstsein als Korrektiv stützen. Diese Untersuchungen dienen mehr illustrierenden Zwecken, da sie nicht zum Test der Hypothese vom Bewusstseins als Korrektiv durchgeführt wurden. Diese Studien liefern aus unserer Sicht jedoch zumindest indirekte Evidenz für diese These.

Beispiel 1: Assoziative und propositionale Verarbeitung von Negationen

Deutsch, Gawronski und Strack (2007) untersuchten in einer Serie von Experimenten die Dissoziation zwischen automatischer, assoziativer (nicht-bewusster) und kontrollierter, propositionaler (bewusster) Verarbeitung von Negationen. Die diese Studie leitende Hypothese bestand darin, dass in einem nicht-bewussten, assoziativen (und damit schnellen) Verarbeitungsmodus eine Negation eines Begriffes nicht verarbeitet wird, sondern stattdessen die semantischen Verbindungen des Begriffes selbst aktiviert werden und dessen subjektive Valenz als positiv oder negativ bestimmen. In einem kontrollierten, propositionalen (und damit langsameren) Verarbeitungsmodus dagegen sollten Negationen sehr wohl als solche verarbeitet werden und damit die Valenz des nicht-negierten Begriffes umdrehen. Ein Beispiel soll dies verdeutlichen: Die Negation von Krieg („kein Krieg“) wird nach Deutsch et al. auf einer assoziativen, automatischen Ebene vornehmlich als „Krieg“ verarbeitet und mit solchen Begriffen wie „grausam“, „Verletzte“, „Tote“ oder „Bombe“ verknüpft. Die subjektive Valenz ist demnach negativ. Erst auf einer kontrollierten, propositionalen Verarbeitungsebene wird auch die Negation verarbeitet, was zu einer Umkehrung der Valenz führt (Krieg = negative Valenz; kein Krieg = keine negative Valenz = positive Valenz). Durch die kontrollierte Verarbeitung kann dieser Hypothese zufolge also eine Korrektur an der automatischen, assoziativen Valenz eines negierten Ausdrucks vorgenommen werden.

Um diese Hypothese zu testen, verwendeten die Autoren das Paradigma des evaluativen Primings. Dies besteht darin, dass die Beurteilung eines Ziel-Wortes (z.B. „Liebe“) hinsichtlich seiner positiven oder negativen Valenz dann schneller ist, wenn die Valenz des kurz zuvor präsentierten Begriffes (so genannte Primes) gleichgerichtet (kongruent) ist, und langsamer, wenn die Valenz inkongruent ist. Wenn der Prime – beispielsweise „kein Krieg“ auf einer assoziativen, automatischen Ebene als „Krieg“ (also negativ valent) verarbeitet wird, dann sollte dies zu einer Verlangsamung der Beurteilung des Ziel-Wortes „Liebe“ als positiv führen. Da die Beurteilung des Ziel-Wortes unmittelbar nach der Darbietung des Primes vorgenommen werden muss, kann in diesem Paradigma nur eine assoziative (schnelle) Verarbeitung stattfinden.

Abb. 5: Reaktionszeiten in der Evaluation von positiv und negativ valenten Zielwörtern nach Darbietung von positiv und negativ valenten nicht-negierten oder negierten Primes (Deutsch, Gawronski und Strack 2006).

Abbildung 5 zeigt die Reaktionszeiten auf die Bewertung des Zielwortes. In der linken Hälfte der Abbildung werden die Bewertungen von Zielwörtern nach Darbietung eines nicht-negierten Primes dargestellt. Hier zeigt sich der allgemein bekannte evaluative Priming-Effekt, der darin besteht, dass ein positives Zielwort (z.B. „Liebe“) schneller beurteilt wird, wenn es auf einen positiven Prime (z.B. „Kuchen“) als auf einen negativen Prime (z.B. „Krieg“) folgt. Auf der rechten Seite von **Abbildung 5** zeigt sich nun, wie vorhergesagt, dass die Negation des Primes keinerlei Effekt auf das Muster der Reaktionszeiten zeigt. Folgt beispielsweise das positiv valente Zielwort „Liebe“ auf den propositional ebenfalls positiv valenten Prime „kein Krieg“, so ist die Reaktionsgeschwindigkeit genauso verlangsamt, als wäre der negativ valente Prime „Krieg“ dargeboten worden. Die Negation scheint also bei einer schnellen, automatischen Verarbeitung keinerlei Rolle für die Valenz eines Reizes zu haben.

Um zu testen, ob bei einer propositionalen, kontrollierten Verarbeitung diese *falsche* automatische Evaluation einer Negation („kein Krieg“ wird ja fälschlicherweise negativ bewertet) *korrigiert* wird, wurde eine weitere Vergleichsbedingung eingeführt. Diese bestand in einer einfachen Beurteilung des Primes ohne nachfolgende Darbietung eines weiteren Zielwortes. Da die Verarbeitung des Primes hier nicht von einem weiteren Stimulus unterbrochen wird, kann eine propositionale Verarbeitung stattfinden. In dieser Bedingung sollte die Negation eines negativ valenten Begriffs auch tatsächlich als positiv beurteilt werden („kein Krieg“ = positiv).

Um die propositionalen Urteile mit den assoziativen Bewertungen der Primes direkt vergleichen zu können, haben die Autoren aus den Reaktionszeitdaten einen Index für den Grad der positiven Bewertung (Positivitäts-Index) berechnet. Der Positivitäts-Index wird für jeden Prime berechnet als die Differenz von Reaktionszeiten auf negative Zielwörter minus den Reaktionszeiten auf positive Zielwörter, die jeweils diesem Prime folgten. Wenn zum Beispiel auf den Prime „kein Krieg“ die Beurteilung positiver Zielwörter als positiv 600 ms gedauert hat, die Beurteilung negativer Wörter als negativ aber nur 550 ms, dann bekommt der Prime „kein Krieg“ einen Positivitäts-Index von -50 ms. Der negative Positivitäts-Index bedeutet hier, dass der Prime negative Urteile gegenüber positiven Urteilen beschleunigt hat, was dafür spricht, dass er selbst automatisch als negativ eingeschätzt wurde.

Abbildung 6 zeigt nun die Positivitäts-Indices für die vier Arten von Primes auf der linken Seite und die evaluativen Urteile für dieselben Primes auf der rechten Seite. Wie erwartet zeigt sich hier eine Dissoziation in der Bewertung. Vergleicht man die Ergebnisse für die Valenzbewertung von Negationen (siehe Hervorhebung in **Abbildung 6**), so zeigt sich eine Umkehrung zwischen den assoziativen und den evaluativen Urteilen. In der propositionalen Verarbeitung wird die Negation eines positiven Begriffs richtigerweise auch negativ bewertet (z.[^]B. „kein Kuchen“ = negativ) und die Negation eines negativen Begriffs positiv (z.[^]B. „kein Krieg“ = positiv). Es kann also insofern von einem Korrektiv gesprochen werden, als bei einer automatischen Verarbeitung (bei assoziativen Urteilen) die Negation keinen Einfluss auf die subjektive Valenz nimmt, während dies bei bewussten, evaluativen Urteilen einer Negation sehr wohl der Fall ist.

Abb. 6: Aus Priming-Effekten berechnete assoziative Positivität der Primes (linker Teil der Abbildung) und deliberative Beurteilung der Primes (rechter Teil der Abbildung). Während die Negation des Primes keinen Einfluss auf die Wirkung der Primes in der assoziativen Beurteilungsaufgabe hat, wirkt sich diese auf die Bewertung in einer propositionalen Verarbeitung deutlich aus (Deutsch, Gawronski und Strack 2006).

Bezogen auf unser Modell kann die Verknüpfung von (negierten und nicht-negierten) Aussagen und Bewertungen als eine Komponente der Verknüpfung von Stimuli und Zielen interpretiert werden – Bewertungen sind keine Ziele, aber sie sind doch eng mit Zielen verbunden: Positive Bewertungen legen ein Annäherungsziel, negative Bewertungen ein Vermeidensziel nahe. Bisher wurden zu der Dissoziation von automatischen, assoziativen im Vergleich zu kontrollierten Reaktionen auf Negationen unseres Wissens nach noch keine weitergehenden Verhaltensstudien durchgeführt, so dass wir über die weiteren Schritte – von Bewertungen über (möglicherweise automatisch aktivierte) Ziele zu Verhaltensweisen – noch nichts wissen. Es liegt aber auf der Hand, dass die Korrektur einer falschen Valenzeinschätzung für eine adäquate Handlungsauswahl essenziell sein kann.

Beispiel 2: Der „false fame“-Effekt

Ein weiteres empirisches Beispiel, das als Evidenz für die These vom Bewusstsein als Korrektiv herangezogen werden kann, ist eine in der Psychologie inzwischen klassische Studie zum so genannten „false fame“-Effekt (Jacoby, Kelley, Brown und Jasechko 1989). Im ersten Teil dieser Untersuchung lasen Versuchsteilnehmer eine Liste von Namen. Im zweiten Teil, der 24 Stunden später stattfand, bekamen sie eine weitere Namensliste, auf der berühmte und unbekannte Personen aufgelistet waren. Ohne dass dies den Versuchsteilnehmenden kenntlich gemacht wurde, waren unter den unbekannteten Personen teilweise Namen der Liste des Vortags wieder aufgenommen worden. Die

Aufgaben bestand nun darin, berühmte Personen auf dieser Liste zu identifizieren. Interessanterweise wurden Namen von vollkommen unbekannt Personen, die am Vortrag bereits einmal gelesen worden waren, überzufällig häufig als „berühmt“ eingeschätzt. Dieser „false fame“-Effekt wird darauf zurückgeführt, dass es ein vages Gefühl der Vertrautheit eines eigentlich unbekannt Namens gibt, wenn man ihn bereits am Vortrag schon einmal gesehen hat. Dieses Gefühl der Vertrautheit wird dann, da Personen sich deren Quelle nicht bewusst sind, als ein Indikator dafür herangezogen, dass es sich um eine berühmte Person handeln müsse. Es handelt sich also um eine Fehlinterpretation der Ursache für das Gefühl der Vertrautheit.

Gestützt wird diese Interpretation des „false fame“-Effektes dadurch – und dies ist im vorliegenden Kontext der eigentlich interessierende Befund dieser Studie –, dass die *fälschliche* Zuschreibung von Berühmtheit dadurch *korrigiert* werden kann, dass den Versuchsteilnehmern mitgeteilt wird, dass sie vielleicht manche der Namen schon einmal auf der Liste des Vortages gelesen haben. Tatsächlich verschwindet der „false fame“-Effekt vollkommen, wenn die Vortagsliste als Quelle des Vertrautheitsgefühls eines Namens auf diese Weise kognitiv verfügbar gemacht wird. Dadurch wird es Personen möglich, die automatische Reaktion auf den internen Reiz „Vertrautheit“ mit dem Urteil der Berühmtheit durch die Herstellung einer neuen Verbindung zu der Quelle der Vertrautheit zu korrigieren.

Beispiel 3: Die Wirkung des Wetters auf die Einschätzung der Lebenszufriedenheit

Die Korrektur eines subjektiven Urteils durch die Bewusstmachung möglicher Einflussgrößen wurde auch in Bezug auf die Einschätzung der eigenen Lebenszufriedenheit in einem sehr einfachen Versuch von Schwarz und Clore (1983) empirisch gezeigt. Leitend für diese Untersuchung war die Frage, wie Personen vorgehen, wenn man ihnen so komplexe Fragen stellt, wie glücklich oder zufrieden sie mit ihrem Leben insgesamt sind. Da die kognitiven Kapazitäten von Menschen nicht ausreichend sind, um alle vergangenen und gegenwärtigen Aspekte ihres Lebens nach ihrer Bedeutsamkeit gewichtet in ein Gesamturteil zu verrechnen, gingen Schwarz und Clore davon aus, dass Menschen ihre Lebenszufriedenheit mit Hilfe von einfachen Heuristiken bestimmen. Eine solche Heuristik ist es, die momentane Stimmung als einen Indikator für die allgemeine Lebenszufriedenheit heranzuziehen. Die momentane Stimmung ist jedoch nicht nur von der allgemeinen Lebenszufriedenheit, sondern auch von einer Vielzahl an situativen Faktoren wie dem Wetter abhängig. Tatsächlich zeigte sich, dass bei einer Telefon-Umfrage, die bei schlechtem Wetter durchgeführt wurde, die Lebenszufriedenheit im Durchschnitt niedriger eingeschätzt wurde als bei gutem Wetter. Der Einfluss des schlechten Wetters auf die Einschätzung der allgemeinen Lebenszufriedenheit konnte jedoch dann rückgängig gemacht werden, wenn Personen das schlechte Wetter vor der Frage nach der Lebenszufriedenheit bewusst gemacht wurde (z.[^]B. mit der Frage „Übrigens, wie ist denn das Wetter heute so in xyz?“). Auch in diesem Fall nehmen Personen eine Korrektur einer direkten Verknüpfung (hier: Wetter – Befindlichkeit), die als Voreinstellung für eine bestimmte Reaktion (hier: Einschätzung der Lebenszufriedenheit) dient, vor, wenn ihnen diese Verknüpfung bewusst wird.

Auch für die Untersuchung der Korrektur von (Fehl-)Urteilen gilt, dass weiterreichendere Studien, die den Zusammenhang von Urteilen mit Zielen und Handlungen untersuchen, noch nicht vorliegen. Zum gegenwärtigen Zeitpunkt bleibt festzuhalten, dass Fehlerurteile beispielsweise im Sinne einer falschen Interpretation eines Gefühls von Vertrautheit (man denke beispielsweise an mögliche Konsequenzen für Zeugenaussagen vor Gericht) oder der Fehleinschätzung der eigenen Lebenszufriedenheit aufgrund zufälliger und transientser Einflüsse durchaus weiterführende Handlungskonsequenzen nach sich ziehen können. Eine Korrektur dieser Fehlerurteile ist für eine angemessene Handlungsauswahl wahrscheinlich von großer Bedeutung.

Diese drei Beispiele psychologischer Studien sind, wie eingangs erwähnt, keine strengen empirischen Belege der in diesem Beitrag aufgestellten These, dass eine Funktion von Bewusstsein darin besteht, als Korrektiv zu fungieren. Auch in Bezug auf das von uns vorgeschlagene hierarchische Handlungsmodell können diese Studien nur als Illustrationen für einige der postulierten Pfade gelten. Eine systematische empirische Untersuchung des Modells steht gegenwärtig noch aus. Ziel dieses Beitrages war es jedoch nicht, ein geschlossenes und empirisch abgesichertes Forschungsprogramm zusammenzufassen. Vielmehr ging es uns darum, eine Möglichkeit zur Konzeptualisierung der Funktion des Bewusstseins aus psychologischer Sicht aufzuzeigen und empirische Forschung hierzu anzuregen. Die experimentellen Methoden, die erfolgreich verwendet wurden, um Dissoziationen zwischen bewussten und unbewussten Prozessen bei der Urteilsbildung aufzuzeigen, sollten sich auch zur Untersuchung analoger Dissoziationen bei der Ziel- und Handlungsauswahl fruchtbar einsetzen lassen.

Bibliographie

- Bargh, John A. (1990): Auto-motives: Preconscious Determinants of Social Interaction. In: Higgins, E. Tory/Sorrentino, Richard M. (Hg.): *Handbook of Motivation and Cognition*. Vol. 2. New York: Guilford, 93-130.
- Bieri, Peter (2001): *Das Handwerk der Freiheit. Über die Entdeckung des eigenen Willens*. München: Hanser.
- Block, Ned (1995): On a Confusion About a Function of Consciousness. *Behavioral and Brain Sciences* (18), 227-247.
- Block, Ned (1981): Psychologism and Behaviorism. *Philosophical Review* (90), 5-43.
- Chalmers, David J. (1996): *The Conscious Mind. In Search of a Fundamental Theory*. New York: Oxford University Press.
- Chaiken, Shelly/Trope, Yaacov (Hg.) (1999): *Dual-Process Theories in Social Psychology*. New York: Guilford.
- Deutsch, Roland/Gawronski, Bertram/Strack, Fritz (2006): At the Boundaries of Automaticity: Negation as Reflective Operation. *Journal of Personality and Social Psychology* (91), 385-405.
- Dijksterhuis, Ap/Nordgren, Loran F. (2006): A Theory of Unconscious Thought. *Perspectives on Psychological Science* (1), 95-109.
- Fitzsimons, Gráinne M./Bargh, John A. (2003): Thinking of You: Nonconscious Pursuit of Interpersonal Goals Associated With Relationship Partners. *Journal of Personality and Social Psychology* (84), 148-163.
- Jacoby, Larry L./ Kelley, Colleen/Brown, Judith/Jasechko, Jennifer (1989): Becoming Famous Overnight: Limits on the Ability to Avoid Unconscious Influences of the Past. *Journal of Personality and Social Psychology* (56), 326-338.
- Kruglanski, Arie W. (1996): Goals as Knowledge Structures. In: Gollwitzer, Peter M./Bargh, John A. (Hg.): *The Psychology of Action: Linking Cognition and Motivation to Behavior*. New York: Guilford Press, 59-68.
- Mayr, Ulrich (2004): Conflict, Consciousness, and Control. *Trends in Cognitive Science* (8), 145-148.
- Merikle, Philip M./Daneman, Meredith (1998): Psychological Investigations of Unconscious Perception. *Journal of Consciousness Studies* (5), 5-18.
- Posner, Michael I./DiGirolamo, Gregory J. (1998): Executive Attention: Conflict, Target Detection and Cognitive Control. In: Parasuraman, Raja (Hg.): *The Attentive Brain*. Cambridge: MIT Press, 401-423.
- Schwarz, Norbert/Clore, Gerald L. (1983): Mood, Misattribution, and Judgments of Well-being: Informative and Directive Functions of Affective States. *Journal of Personality and Social Psychology* (45), 531-523.