

Andreas Kemmerling

Selbstkenntnis als ein Test für den naturalistischen Repräsentationalismus

Selbstkenntnis interessiert Philosophen aus vielfältigen Gründen. Mich interessiert sie im folgenden als ein Probestein für gewisse philosophische Auffassungen über Intentionalität. Die leitende Frage meiner nachfolgenden Betrachtungen wird sein: Wie kommen naturalistisch-repräsentationalistische Konzeptionen von Intentionalität mit dem Phänomen der Selbstkenntnis zurecht?

Nicht das Wesen, der Umfang oder der Wert von Selbstkenntnis sind also hier das Thema, sondern die Frage, ob man unter gewissen Voraussetzungen, die in der zeitgenössischen Philosophie des Geistes von vielen gemacht werden, bestimmten Aspekten der Selbstkenntnis gerecht werden kann. Dabei werde ich mich auf solche Aspekte von Selbstkenntnis beschränken, die meines Erachtens unbestreitbar sind. (Unbestreitbar jedenfalls insofern, als diese Aspekte von Selbstkenntnis bei normalen Personen unter normalen Umständen immer anzutreffen sind.) Und ich werde mich am Ende auf diejenige Konzeption konzentrieren, die mir unter den derzeit vertretenen naturalistisch-repräsentationalistischen Theorieansätzen noch am aussichtsreichsten erscheint, und zwar ist das die von Fred Dretske.

Was Selbstkenntnis sein soll

Unter Selbstkenntnis kann man allerlei verstehen. In der Umgangssprache gibt es so etwas anscheinend nicht; jedenfalls findet sich im sechsbändigen *Großen Wörterbuch der deutschen Sprache* aus dem Jahre 1980 kein Eintrag zwischen »Selbstkasteiung« und »Selbstklebefolie«.¹ Auch in der deutschen Philosophen-

¹ *Das Große Wörterbuch der deutschen Sprache* (hg. und bearbeitet vom Wiss. Rat und den Mitarbeitern der Dudenredaktion unter Leitung von Günther Drosdowski), Mannheim/Wien/Zürich 1980, Bd. 5, S. 2375.

Fachsprache scheint keine eindeutige Regelung vorgegeben zu sein. Deshalb möchte ich deutlich machen, wie ich diesen Terminus verwende.

Die Selbstkenntnis einer Person umfaßt alles, was die Person über sich selbst weiß – genauer gesagt: was sie über sich selbst *als sich selbst* weiß. Der Zusatz »als sich selbst« soll hervorheben, daß Selbstkenntnis sog. De-se-Wissen ist; bloßes De-re-Wissen über die eigene Person ist noch keine Selbstkenntnis. Ein Beispiel: Jemand mag wissen, daß der Mann, den er im Spiegel von hinten sieht, ein großes Kreidekreuz auf dem Rücken hat; da er selbst dieser Mann ist, den er im Spiegel von hinten sieht, weiß er also (wie es heißt: »de re«) von sich, daß er ein großes Kreidekreuz auf dem Rücken hat; aber es kann dennoch sein, daß er dies nicht von sich *als sich selbst* weiß. Er könnte ja im selben Moment den Satz »Ich habe ein Kreidekreuz auf dem Rücken« für unsicher oder gar falsch halten. Dann wäre sein Wissen kein De-se-Wissen und gehörte mithin nicht zu seiner Selbstkenntnis. Die Möglichkeit von De-re-Wissen über sich selbst, das kein De-se-Wissen ist, werde ich von nun an beiseite lassen; deshalb wird es auch nicht mehr nötig sein, den erwähnten Zusatz zu machen. Wenn ich im folgenden davon spreche, was eine Person über sich selbst weiß, dann meine ich damit immer ihr De-se-Wissen. Eine weitere terminologische Festsetzung vorab: Alles, was eine Person an Nicht-Kontingentem über sich weiß, soll hier ebenfalls unberücksichtigt bleiben. Anders gesagt: Wenn E eine Eigenschaft ist, die einer Person schon aus rein logischen oder begrifflichen Gründen zukommt, dann gehört das, was sie mit dem Satz »Ich habe E« sagt, nicht zu ihrer Selbstkenntnis. (Beispiele für das, was auf diese Weise ausgeschlossen wird, sind: wissen, daß man, falls man reich ist, reich ist; wissen, daß man keine Primzahl ist; und so weiter.)

Unter all dem, was jemand über sich selbst weiß, gibt es einiges, das er in bemerkenswert besonderer Weise weiß. Ein Beispiel: Wer weiß, daß er jetzt gerade an seine Bankschulden denkt, der weiß dies in bemerkenswert besonderer Weise – in einer Weise, die sich erheblich von seinem Wissen unterscheidet, daß er Bankschulden hat. Beiderlei Wissen ist zur Selbstkenntnis zu rechnen, so wie ich diesen Begriff zu verwenden vorschlage; aber das Wissen um die eigenen Bankschulden ist kein besonderes. Die relevanten Besonderheiten, die ich im Auge habe, werden in der

Literatur mit Ausdrücken wie den folgenden hervorgehoben: »introspektiv«, »unmittelbar«, »a priori«, »basal«, »mit Erster-Person-Autorität«, »dank privilegiertem Zugang«, »unfehlbar«, »nicht korrigierbar« und ähnliches mehr. Alle diese Charakterisierungen sind erläuterungsbedürftig; manche sind meines Erachtens durchaus irreführend (zum Beispiel »introspektiv«, »a priori«, »basal« und »dank privilegiertem Zugang«). Dennoch will ich an dieser Stelle nicht versuchen, eine exakte Erläuterung dafür zu geben, worin jene besondere Selbstkenntnis besteht, sondern schlicht unterstellen, daß die beiden folgenden Arten von Wissen in jene Sonderklasse der Selbstkenntnis fallen:

(a) Wissen, das man darüber hat, was man im gegenwärtigen Moment denkt oder glaubt;

(b) Wissen, das man darüber hat, wie einem im gegenwärtigen Moment das vorkommt, was man wahrzunehmen meint.

Von Wissen des Typs (a) ist in Sätzen wie den folgenden die Rede: »Harvey weiß, daß er jetzt den Gedanken hat, daß es regnet«; »Harvey weiß, daß er jetzt glaubt, daß es schneit«; Wissen vom Typ (b) liegt vor, wenn ein Satz wie der folgende wahr ist: »Harvey weiß, daß ihm die Flüssigkeit im Mund jetzt wie ein junger Merlot vorkommt«, »Harvey weiß, daß ihm das, worauf er blickt, jetzt blau vorkommt«. – Jene durch (a) und (b) exemplifizierte Sonderklasse der Selbstkenntnis dürfte mit gewissem Recht durch das Anhängsel »à la Descartes« etikettiert werden.

Descartes hatte die Gewißheit, die ein Denker von seinen eigenen momentanen Gedanken hat, zum Ausgangspunkt seiner konstruktiven Überlegungen gemacht, mit denen er nachweisen wollte, daß und wodurch es fehlbaren Menschen möglich ist, zu echtem Wissen zu gelangen. Allerdings hat er, soweit ich sehe, derartige Gewißheiten selbst (wie »Ich denke jetzt, daß ich existiere« oder »Ich denke jetzt, daß alles zweifelhaft ist«) nie als echtes Wissen (*scientia*) bezeichnet. Er zieht es beim Thema Selbstkenntnis vor, andere Wörter als *scire* und *scientia* zu verwenden: Wir sind uns unserer Gedanken *bewußt* (*consciūs*; zum Beispiel AT VII 246);² wir haben *Bekanntschaft* (*notitia*, AT VII 357) mit uns selbst; wir können uns im Schlaf dessen *gewahr* (*advertere*, AT VII 359) sein, daß wir träumen; im Denken *erle-*

² Descartes' Schriften werden nach der Ausgabe von Charles Adam und Paul Tannery (*Œuvres de Descartes*, Paris 1897-1910) zitiert.

ben wir (*experiri*, AT VII 358, 427), daß wir denken; und so weiter.

In diesem Zusammenhang nicht von Wissen zu sprechen (und übrigens natürlich auch nicht von Glauben), halte auch ich für angemessen. Wer denkt, daß es regnet, von dem kann man nicht sagen, er wisse oder glaube währenddessen, daß er dies denkt. Genausowenig läßt sich von ihm sagen, er wisse oder glaube es währenddessen nicht; und selbst die negationshervorziehende Redeweise: es sei währenddessen nicht der Fall, daß er es wisse oder glaube, ist unpassend. Das gesamte Glaubens- und Wissensidiom ist hier fehl am Platz.

Selbstkenntnis – oder »Selbstwissen«, wie manche sagen³ – à la Descartes ist demnach streng genommen gar kein Wissen. Doch wie schief diese Redeweise auch ist, sie hat sich im philosophischen Sprachgebrauch festgesetzt. Das bedaure ich, aber ich werde in der vorliegenden Arbeit keine sachlich passendere Terminologie vorschlagen, sondern wider besseres Wissen so reden, als handele es sich wirklich um eine Gattung von Wissen. Meine Rechtfertigung für diese terminologische Mitläuferei besteht darin, daß ich meinem Gegner – jedem, der Intentionalität repräsentationalistisch naturalisieren möchte – Einwände zu einem ganz andern Punkt machen möchte, den er nicht als Wortklaubelei abtun kann.

Im folgenden geht es also um Selbstkenntnis à la Descartes: kontingentes De-se-Wissen, das von den eigenen momentanen Gedanken, Überzeugungen und Wahrnehmungseindrücken handelt (und von sonst nichts). Irrtum und sogar Zweifel sind dabei für den Betreffenden ausgeschlossen; und zwar in folgendem Sinn: Wer zu irgendeinem Moment den Gedanken, Glauben oder Wahrnehmungseindruck hat, daß es regnet, der kann darüber im selben Moment nicht irren; er kann dann nicht einmal – sozusagen im selben Atemzug – bezweifeln, daß er in diesem Moment denkt, glaubt oder den Eindruck hat, daß es regnet.

Daß es angesichts all dessen verlockend ist, von Wissen zu sprechen, liegt auf der Hand. Denn wenn jemand den Gedanken hat, daß es regnet, dann gibt es etwas Kontingentes, das im selben Moment wahr ist, ihm nicht entgeht, über das er sich nicht irren

3 Auch dieser Terminus hat es noch nicht zu einen Eintrag in das *Große Wörterbuch* gebracht.

und das er nicht bezweifeln könnte: und zwar eben just dies, daß er im betreffenden Moment diesen Gedanken hat. Und es liegt nahe (obwohl es falsch ist), etwas als Wissen zu bezeichnen, das all diese außerordentlichen Eigenschaften hat.

Da es im folgenden ausschließlich um Selbstkenntnis à la Descartes geht, werde ich – der Einfachheit halber und weil keine Verwechslungen mit andern Arten der Selbstkenntnis drohen – den Zusatz »à la Descartes«, so schön er ist, weglassen. Ich werde »Selbstkenntnis« sagen, aber Selbstkenntnis à la Descartes meinen.

Was Naturalisierung von Intentionalität sein soll

Unter der sog. »Naturalisierung« von sog. »Intentionalität« mag man nach Durchsicht der einschlägigen Monographien, Zeitschriften und Sammelbände der letzten dreißig Jahre allerlei verstehen. Was sich in der Literatur zu diesen Stichwörtern findet, ist sehr verschiedenartig. Entsprechend ist auch der Terminus »Naturalisierung von Intentionalität« kein Etikett eines wohlgeführten Markenartikels. Wer diesen Terminus verwendet, muß sagen, was er darunter versteht.

So sei es. Zunächst einmal: Es ist ein Projekt, keine Position. Eher ein architektonischer Plan zur Ausgestaltung eines in Umrissen vorgezeichneten gastfreundlichen Hauses. Weniger ein feldherrlicher Plan zur Verteidigung einer bereits errichteten trutzhaften Festung. Das gastfreundliche Haus, das sind die begrifflichen Möglichkeiten der Naturwissenschaften. Wer sich diesem Projekt zurechnet, sucht nach irgendeiner würdigen Einbürgerung der sog. intentionalen Phänomene in die Welt der sogenannten natürlichen Phänomene. Was ist das Erfolgskriterium? Wann wäre solch ein Eingliederungsversuch als gelungen zu betrachten? Reichte es, wenn die intentionalen Phänomene so analysiert würden, daß sie erkennbar in den Rahmen unseres derzeitigen naturwissenschaftlichen Weltbildes hineinpassen? Oder müßte etwas sehr viel Anspruchsvolleres gezeigt werden, etwa daß eine naturwissenschaftliche Theorie dieser Phänomene möglich ist? All das ist nach meinem Eindruck einigermaßen unklar. »Das« Projekt der Naturalisierung von Intentionalität ist reichlich verschwommen.

Wichtig und soweit ich sehe unumstritten ist, daß eine Naturalisierung keine Reduktion verlangt. Angenommen es gelänge, für beliebige intentionale Phänomene I in nicht-intentionalistischer und naturwissenschaftlich akzeptabler Terminologie eine Eigenschaft B (die neuronalen oder andern geeigneten Hirn-Zuständen zukommen kann) anzugeben, so daß folgendes gilt: Der Umstand, daß jemand sich in einem neuronalen Zustand mit der Eigenschaft B befindet, ist eine nichttriviale hinreichende Bedingung dafür, daß er I hat. Und weiterhin angenommen, daß diese Bedingung in fast jedem tatsächlich auftretenden Fall erfüllt ist, in dem jemand I hat. Wenn so etwas gelänge, dann wäre das intentionale Phänomen zwar naturalisiert, aber nicht physikalistisch reduziert.

Eine derartige Konzeption der Naturalisierung von Intentionalität mag J. Fodor einmal vorgeschwebt haben.⁴ Doch auch sie halte ich, trotz ihres nicht-reduktionistischen Charakters, immer noch für überzogen. Es gibt jede Menge intentionaler Phänomene, deren Existenz zwar unbezweifelbar ist, die aber nicht in dieser Weise naturalisierbar sind. Zu der geschilderten Konzeption gehört ja die Annahme, es müsse sich ein Hirnzustand finden lassen, der als Träger derjenigen (nicht-intentional beschreibbaren) Eigenschaft fungiert, durch die die intentionale Eigenschaft als naturalisiert zu betrachten ist. Diese Annahme mag im Falle episodischer oder sehr kurzfristiger intentionaler Phänomene (wie Einen-Einfall-Haben, Einen-Einwand-Durchdenken usw.) plausibel erscheinen. Im Hinblick auf einige lang andauernde, dispositionale intentionale Phänomene ist sie von vornherein verfehlt. Man denke nur an solche Überzeugungen, die einer zwar über einen langen Zeitraum hinweg hat, die aber von einem Sachverhalt handeln, an den er noch nie einen Gedanken verschwendet hat. Ein Beispiel: Seit vielen, vielen Jahren glaube ich, daß die Welt schon vor dem 22. Februar 1679 existiert hat. Jeder normale Mensch hat unendlich viele solcher Überzeugungen, und natürlich nicht nur Überzeugungen, sondern auch Wünsche, Präferenzen und so weiter. Die Annahme, für jede einzelne von ihnen müsse sich ein Hirnzustand im Subjekt finden lassen, der dank einer besonderen Eigenschaft, die ihm zu-

4 Vgl. Jerry Fodor, *A Theory of Content and Other Essays*, Cambridge, Mass. 1990, Kapitel 3, insb. S. 96.

kommt, das Subjekt zu einem macht, das die entsprechende Überzeugung hat – diese Annahme ist allein schon angesichts der Anzahl der Zustände, um die es geht, inakzeptabel.

Wollte man sich nun im Projekt der Naturalisierung auf solche intentionalen Phänomene beschränken, in bezug auf die man sich sinnvollerweise geeignete Träger-Zustände im Hirn erwarten kann, dann bliebe allzu vieles aus dem Bereich des Intentionalen bei der Naturalisierung ausgeschlossen. Fodor ist angesichts dieses Problems willens, sich auf sogenannte »tatsächlich auftretende« Überzeugungen [*occurrent beliefs*] als die Kernfälle [*core cases*] seiner naturalistischen Theorie zu beschränken.⁵ Doch das ist nicht akzeptabel, solange es um die Naturalisierung von Überzeugungen geht. Denn das, was Fodor an der zitierten Stelle als »tatsächlich auftretende« Überzeugungen bezeichnet, sind zwar intentionale Phänomene, aber keine Überzeugungen. Vielmehr handelt es sich bei ihnen um episodische Gedanken, für die gilt: ihr Inhalt wird vom Subjekt in diesem Moment geglaubt; das heißt, es sind Gedanken, zu denen das Subjekt auch eine inhaltlich entsprechende Überzeugung hat. Diese Episoden sind keine Überzeugungen, sondern setzen deren Vorhandensein voraus. Wollte man sich bei der Naturalisierung auf derartige Gedankenepisoden beschränken, dann klammerte man gerade etwas aus, was von allen Seiten als grundlegendes intentionales Phänomen betrachtet wird: Überzeugungen.

Die Begriffsverwirrungen, die Fodor bei dieser Gelegenheit anrichtet, sind aufschlußreich. Er kontrastiert eine »manifeste Überzeugung« – etwas Episodisches, das in Wirklichkeit gar keine Überzeugung ist – mit einer »bloß dispositionalen« Überzeugung. Er übergeht das simple Faktum, daß jede Überzeugung (besser: jedes Überzeugtsein) ein dispositionaler Zustand ist. Echte Überzeugungen werden durch diese terminologische Falschmünzerei zunächst einmal in eine abgelegene Ecke geschoben, metaphysisch gesehen sind sie offenbar zweite Wahl: eben »bloß« dispositional. Mittels eines weiteren terminologischen Kunstgriffs werden sie dann endgültig aus dem Bereich der tatsächlich vorhandenen Geisteszustände hinausgedrängt. Fodor setzt sie unter der Hand mit *potentiellen* Überzeugungen gleich.

⁵ Siehe dazu Jerry Fodor, *Psychosemantics*, Cambridge, Mass. 1987, S. 21-25.

Dies sind gemäß Fodors Terminologie solche Überzeugungen, die jemand im betreffenden Augenblick gar nicht hat, aber dadurch erwerben könnte, daß er allein aus den Überzeugungen, die er tatsächlich hat, Schlüsse zieht, die von ihm akzeptierten Schlußmustern genügen. Ein Beispiel zur Verdeutlichung: Da sei jemand, der jetzt zwar nicht glaubt, daß es unendlich viele Primzahlen gibt, der aber durch einen Überlegungsgang, in dem er nur akzeptiert, was er bisher schon akzeptiert, zu der Einsicht gelangen könnte, daß es unendlich viele Primzahlen gibt; so einer glaubt *potentiell*, daß es unendlich viele Primzahlen gibt. Man beachte: Was zu einem Zeitpunkt potentiell geglaubt wird, wird zu diesem Zeitpunkt nach Voraussetzung gar nicht geglaubt. Kurz, potentielle Überzeugungen sind solche, die man zwar nicht hat, aber »leicht« haben könnte.

Auf solchen Wegen – von »bloß dispositional« über »ohne eine kausale Rolle in aktualen Geistesprozessen« zu »potentiell« – gelangt Fodor dazu, Überzeugungen auszublenden. Er verwendet zwar das Wort, aber redet von etwas anderm. Damit steht er nicht allein; auch Lycan betreibt die Naturalisierung von dem, was er Überzeugungen nennt, unter ähnlichen Umetikettierungen. Auch er spricht von einer »Kernklasse von manifesten [*occurrent*] Überzeugungen«, und zwar seien dies »tatsächliche episodische Zustände von Personen (wiewohl nicht unbedingt Bewußtseinszustände)«. ⁶ Wiederum ist der angebliche zentrale Fall des zu naturalisierenden Phänomens überhaupt kein Exemplar dieses Phänomens.

Es geschieht also gelegentlich, daß das, was da unter großem Getöse naturalisiert werden soll, eine bei genauerem Hinsehen nur noch sehr kleine Intentionalität ist. Und die ist nicht das, als was sie ausgegeben wird. Überzeugungen sind eben, im Gegensatz zu dem, als was Lycan sie ausgibt, etwas wesentlich anderes als dies: »Urteile – Episoden geistigen Zustimmens, wenn man so will«. ⁷ Aber warum will man denn das überhaupt: Episoden des Urteilens an die Stelle von Überzeugungen bugsieren? Woran liegt es, daß die zentralen intentionalen Phänomene (wie Überzeugungen zum Beispiel) schon im terminologischen Vorgeplänkel durch etwas anderes ersetzt werden?

⁶ William G. Lycan, *Judgement and Justification*, Cambridge 1988, Kapitel 1-3, Zitate S. 6 und 15.

⁷ Ebd., S. 6.

Nach meinem Eindruck liegt dies an einer Reihe von Vorurteilen, die allzu viele Philosophen zur Zeit über einige Dinge haben und die nicht gut zu der Natur intentionaler *Zustände* passen. Eines dieser Vorurteile – ein sehr komplexes – ist metaphysischer Art: Geistige Phänomene sind nur dann wirklich vorhanden, wenn sie in einem Subjekt »manifeste« Vorkommnisse [*tokens*] haben, die »kausal aktiv« werden können. Ein zweites Vorurteil besagt: Intentionale Geistesphänomene sind repräsentational und involvieren also Repräsentationen. Im Lichte solcher Unterstellungen fällt es schwer, damit zurechtzukommen, daß Überzeugungen eben gar keine Zustände sind, auf die eine *type/token*-Unterscheidung anwendbar ist,⁸ und vielleicht nicht einmal Zustände, die überhaupt Repräsentationen involvieren.

Die Naturalisierung von Intentionalität ist ein Projekt, das – wie mir scheint – von fast jedem, der sich ihm zurechnet,⁹ unter wenigstens einer externen Prämisse betrieben wird, die ich für falsch halte. Eine solche von außen gerne herangetragene Prämisse ist die von der Wahrheit der Token-Identitätstheorie; eine andere ist die von der Wahrheit des Repräsentationalismus. Doch keine von beiden paßt auf intentionale Zustände wie Überzeugungen. Dennoch gibt es Überzeugungen – und das heißt ja nichts anderes als: Es gibt Lebewesen, die Überzeugungen haben. Diese Tatsache paßt trefflich ins naturwissenschaftliche Weltbild: nicht schlechter jedenfalls als die Tatsache, daß manchmal Vollmond ist. Der sog. Eliminativismus, dem zufolge es intentionale Zustände wie Überzeugungen in Wahrheit nicht gibt, ist eine naturalistische Reaktion auf die Einsicht in die Unvereinbarkeit von Token-Identitätstheorie, Repräsentationalismus und Intentionalitätsrealismus. Dies ist jedoch eine Fehlreaktion. Der Eliminativist ist ein Naturalist, der Intentionalität zwar ganz gerne naturalisiert sähe, sich aber vorweg auf Gedeih und Verderb andern Auffassungen als dem Naturalismus verschrieben hat. Er fühlt sich gedrängt, Überzeugungen zu eliminieren, weil

8 Siehe dazu meine beiden Arbeiten: »Zur sog. Naturalisierung von Intentionalität«, in: Axel Burri (Hg.), *Sprache und Denken*, Berlin/New York 1997, und »Überzeugungen für Naturalisten«, in: Georg Meggle (Hg.), *Analyomen 2*, Berlin/New York 1997.

9 Ich spreche von »fast jedem« nicht zuletzt deshalb, weil ich mich selbst diesem Projekt zurechne.

er die starke Überzeugung hat, die Token-Identitätstheorie müsse wahr sein und der Repräsentationalismus noch dazu.

Man muß also, so denke ich, unterscheiden zwischen dem eigentlichen Anliegen, Intentionalität zu naturalisieren, auf der einen Seite und irgendwelchen hinzukommenden Vorurteilen und Unterstellungen, die sich in sei's auch noch so vielen gutgemeinten Projektskizzen finden, auf der andern Seite. Es fragt sich: Was ist hier Kern, was verzichtbare Zutat?

Was Quine einmal als »Ablehnung des Geistes als einer zweiten Substanz, die zum Körper noch hinzukommt« bezeichnet hat,¹⁰ ist in meinen Augen der springende Punkt des Naturalismus. Der Naturalist erkennt keine konkreten Einzeldinge in der Welt an, die zu den materiellen (von »der besten physikalischen Theorie« postulierten) noch hinzukämen. Es steht ihm jedoch frei, gewisse nicht-materielle Eigenschaften anzuerkennen, das heißt solche, die sich nicht auf physikalische Eigenschaften zurückführen lassen. Der Naturalist kann zwar nicht beliebige nicht-materielle Eigenschaften anerkennen, zum Beispiel nicht die Eigenschaft, ein immaterieller Geist zu sein. Aber er kann jede nicht-materielle Eigenschaft zulassen, die folgende Bedingungen erfüllt: (a) sie superveniert auf der Gesamtheit der physikalischen Eigenschaften und ihrer Verteilung, (b) sie kommt materiellen Einzeldingen zu (oder Entitäten, die logische Konstrukte aus naturalistisch akzeptablen Elementen sind), und (c) sie ist durch ein in seiner Anwendung intersubjektiv geregeltes Prädikat ausdrückbar, das erklärende Kraft hat. Die Bedingungen (a)-(c) reichen hin, um eine Eigenschaft naturalistisch respektabel zu machen, selbst wenn sie sich nicht auf physikalische Eigenschaften zurückführen läßt, das heißt auf diejenigen Eigenschaften, von denen in den grundlegenden Gesetzen der Physik die Rede ist.

Mit einer Unterscheidung Quines läßt sich dieser Punkt so formulieren: Der Naturalist ist Physikalist, was seine Ontologie angeht: Alle konkreten Einzeldinge, von denen seine Ganz Große Theorie (das heißt die Theorie, die alle kontingenten Tatsachen erfaßt) handelt, sind materielle Gegenstände. Aber der Naturalist ist nicht auf den Physikalismus festgelegt, was die Ideologie angeht: Nicht alle Grundprädikate seiner erträumten Ganz Großen Theorie drücken physikalische Eigenschaften aus. Die Trajekto-

10 Willard V. Quine, *Quiddities*, Cambridge, Mass. 1987, S. 133.

rien, Ladungszustände, usw. aller physikalischen Gegenstände (einschließlich derjenigen von menschlichen Körpern und ihren Teilen) sind aus ihren physikalischen Eigenschaften zu erklären. Nicht so das Benehmen, die Handlungen, Gedanken, Gefühle und intentionalen Zustände von Subjekten – obgleich auch Subjekte physikalische Gegenstände sind. Das Benehmen, Handeln, Denken, Fühlen usw. ist physikalisch nicht zu erklären, wohl aber mit nicht-materiellen Eigenschaften, die in irreduzibler Manier auf den physikalischen aufsitzen. (Dies ist natürlich nicht Quines Auffassung von Naturalismus; von ihm übernehme ich hier nur die Unterscheidung zwischen der Ontologie und der Ideologie einer Theorie: das heißt zwischen den Entitäten, über die in den Sätzen der Theorie quantifiziert wird, und den Prädikaten, die in den Sätzen der Theorie auftauchen.¹¹)

Nach dieser Auffassung davon, was Naturalismus ist, muß das Projekt der Naturalisierung von Intentionalität nicht notwendigerweise ein Projekt der Reduktion des Intentionalen auf das Physikalische sein. Intentionalität wäre auch dann schon naturalisiert, wenn sich folgendes zeigen ließe:

(a) Die intentionalen Eigenschaften von Individuen und ihre Verteilung in der Welt supervenieren auf der Welt-Gesamtheit der physikalischen Eigenschaften und deren Verteilung in der Welt; mithin exemplifizieren in physikalisch vollständig gleichartigen Welten die gleichen Entitäten dieselben intentionalen Eigenschaften.

(b) Intentionale Eigenschaften kommen materiellen Einzeldingen zu – oder solchen Entitäten, die logische Konstrukte aus naturalistisch akzeptablen Elementen sind. (Mit dem Zusatz habe ich folgendes im Sinn: Angenommen, manche intentionalen Eigenschaften kommen nur Personen zu; und außerdem angenommen, Personen sind nicht einfache materielle Einzeldinge; dann gibt es wenigstens drei Möglichkeiten: (i) Personen sind nicht-materielle Einzeldinge sui generis; (ii) Personen sind materielle-Einzeldinge-qua-Träger-von-Eigenschaften-des-Typs-X, wobei X ein naturalistisch inakzeptabler Eigenschaftstyp ist,

11 Zur Unterscheidung siehe Willard V. Quine, »Ontology and Ideology«, in: *Philosophical Studies* 2, 1951, 11-15. Zu Quines Auffassung von Naturalismus siehe insbesondere »Things and Their Place in Theories«, in: ders. *Theories and Things*, Cambridge/London 1981, 1-23, sowie seinen Beitrag in diesem Band.

oder (iii) Personen sind materielle-Einzeldinge-qua-Träger-von-Eigenschaften-des-Typs-Y, wobei Y ein naturalistisch akzeptabler Eigenschaftstyp ist. Der Zusatz soll (iii) zulassen und die andern beiden Möglichkeiten ausschließen.)

(c) Intentionale Eigenschaften sind durch in ihrer Anwendung intersubjektiv geregelte Prädikate ausdrückbar, die erklärende Kraft haben.

(d) Alle intentionalen Phänomene lassen sich auf intentionale Eigenschaften zurückführen; anders gesagt: zum Bereich des Intentionalen gehören keine irreduziblen konkreten Einzeldinge intentionaler Art.

Mit dem Nachweis, daß die Feststellungen (a)-(d) zutreffen, wäre die Naturalisierung von Intentionalität vollbracht. Für die Ausführung dieses speziellen Projekts – einen Nachweis für (a)-(d) zu führen – sehe ich Chancen; und da ich alle vier Behauptungen für wahr halte, halte ich Intentionalität für ein natürliches Phänomen.

Die philosophischen Positionen, mit denen ich mich anschließend auseinandersetzen werde, propagieren vehementere Naturalisierungsprojekte, in denen wenigstens die sog. Token-Identitätstheorie als wahr vorausgesetzt wird. Diese Voraussetzung läuft, grob gesagt, auf folgendes hinaus: Wann immer ein Individuum *a* eine intentionale Eigenschaft *I* exemplifiziert, exemplifiziert *a* ein physikalisch beschreibbares Merkmal *P*, so daß gilt: as *I*-Exemplifikation ist zum betreffenden Zeitpunkt identisch mit as *P*-Exemplifikation. (Anders gesagt: Zum betreffenden Zeitpunkt exemplifiziert *a* *I* dadurch, daß *a* *P* exemplifiziert.) Das klingt ungebührlich kompliziert. Mit einem Beispiel läßt sich vielleicht besser verstehen, was das eigentlich soll. Nehmen wir an, daß Harvey jetzt glaubt, daß es regnet. In den Jargon übertragen: Er exemplifiziert zum jetzigen Zeitpunkt die intentionale Eigenschaft des Glaubens-daß-es-regnet. Der Anhänger der Token-Identitätstheorie behauptet nun folgendes: Dann befindet sich Harvey jetzt auch in irgendeinem bestimmten Neuro-Zustand, zum Beispiel dem Zustand, daß bestimmte neuronale Verbindungen im corticalen Areal Nr. So-und-so die-und-die Aktivitäten aufweisen; nennen wir diesen Neuro-Zustand »Zoppel«. Im Jargon: Harvey exemplifiziert zum jetzigen Zeitpunkt die physikalische Eigenschaft des Zoppelns. Und worauf es bei alle-

dem ankommt, ist einfach, daß dann gilt: Das Zoppeln in Harvey ist sein Glauben-daß-es-regnet.¹²

Jede derartige Theorie ist im Hinblick auf intentionale Zustände falsch.¹³ Auf eher ereignishafte oder vorgangsartige intentionale Phänomene (wie zum Beispiel plötzliche Einfälle oder Schritt für Schritt sich entwickelnde Überlegungen) mag sie besser passen. Doch sind intentionale Zustände wie das Glauben, das Wollen oder das Bevorzugen grundlegende intentionale Phänomene. Sie lassen sich auf keine anderen intentionalen Phänomene zurückführen, die nicht von der Art eines Zustands wären. Was es heißt zu glauben, daß es regnet, das ist nicht zurückführbar auf intentionale Ereignisse.

Es gibt bekanntermaßen Anhänger der Token-Identitätstheorie, die trotzdem zu allem bereit sind. Sie nennen sich Eliminativisten. Sie sind entschlossen, auf Biegen und Brechen an folgender These festzuhalten: »Ein intentionales Phänomen, das in keiner Token-Identitätstheorie einzufangen ist, existiert eben gar nicht. Basta.« Eliminativisten sind bereit, den Schluß zu ziehen, daß es intentionale Zustände wie das Glauben nicht gibt. Manchmal behaupten sie sogar, ihn gezogen zu haben. Wer diesen Schluß zu ziehen bereit ist, ist bereit anzunehmen, daß es intentionale Zustände nicht gibt. Diese Bereitschaft verlangt Stärke. Man müßte zum Beispiel bereit sein anzunehmen, daß es ein Annehmen eigentlich nicht gibt. Das vermag nicht jeder. – Lassen wir das.

Was Repräsentationalismus hier heißen soll

Unter Repräsentationalismus kann man allerlei verstehen. Hier meine ich damit die Auffassung, daß alle intentionalen Phänomene des Geistigen repräsentational sind. Das soll heißen, daß ein Repräsentationalist, so wie ich diese Bezeichnung hier verwende,

¹² Der Vorspann: »In diesem Moment gilt« ist nicht überflüssig. Der Anhänger der Token-Identitätstheorie läßt ausdrücklich die Möglichkeit zu, daß Harvey bei anderer Gelegenheit zwar wiederum glaubt, daß es regnet, bei dieser Gelegenheit aber nicht zoppelt. Ein anderer Neuro-Zustand muß dann an Zoppels Stelle treten.

¹³ Siehe dazu meine beiden Arbeiten, die ich in Anmerkung 8 erwähnt habe.

folgendes behauptet: Wann immer ein intentionales Geistesphänomen I mit dem Inhalt x (also zum Beispiel die Überzeugung, daß es regnet, der Wunsch, es möge schneien, usw.) bei einem System S vorliegt, gibt es in S eine mentale Repräsentation m derart, daß m x zum Inhalt hat. Dabei sind mentale Repräsentationen irgendwelche konkreten Hirnstrukturen, deren naturalistische Dignität außer Frage steht. Der naturalistische Repräsentationalist sieht sein Hauptproblem gerne darin, eine Theorie darüber zu entwickeln, wie eine mentale Repräsentation (also eine Neurostruktur irgendeiner Art) überhaupt einen Inhalt bzw. gerade den Inhalt hat, den sie hat. Die Theorie, die er anstrebt, soll es erlauben, eine Feststellung wie zum Beispiel

Die Neurostruktur n hat den Inhalt, daß es regnet

so zu analysieren, daß im Analysans kein semantischer oder intentionaler Begriff vorkommt. Er möchte beispielsweise den Sachverhalt, daß Harvey glaubt, daß es regnet, folgendermaßen naturalistisch analysieren: In Harveys Hirn gibt es ein konkretes Vorkommnis n des Neurostrukturtyps N, und n bedeutet, daß es regnet, und Harvey steht in der naturalistisch beschreibbaren Beziehung B zu n. Ferner gilt: Wenn Harvey in B zu einem beliebigen Vorkommnis einer Neurostruktur des Typs N steht, dann glaubt Harvey das, was jenes Vorkommnis bedeutet.

Das am Repräsentationalismus, worum es mir hier geht, betrifft keine Einzelheiten des Stils oder Formats der Repräsentation. Ob bildlich oder sprachlich, analog oder digital, parallel oder seriell, symbolisch oder konnektionistisch, amorph oder strukturiert – all das ist nebensächlich für die nachfolgenden Betrachtungen. Repräsentationalist ist, wer dieses Schlußschema als gültig unterstellt:

Das System S weist das intentionale Phänomen I mit dem Inhalt x auf.

Es gibt in S eine konkrete Repräsentation m derart, daß m x zum Inhalt hat.

Mehr soll mit dem Terminus »Repräsentationalismus« im folgenden nicht gemeint sein.

Naturalistisch-repräsentationalistische Modelle der Selbstkenntnis

Das Wissen, das man hat, wenn man weiß, daß man die-und-die Überzeugung hat, ist ein intentionaler Zustand, der von einem intentionalen Zustand handelt. Wer weiß, daß er glaubt, daß es regnet, der weiß etwas (intentionaler Zustand #1), und sein Wissen handelt davon, daß er die Überzeugung hat, daß es regnet (intentionaler Zustand #2).

Nennen wir einen intentionalen Zustand, der von einem nicht-intentionalen Sachverhalt handelt, einen intentionalen Zustand erster Stufe, und nennen wir einen intentionalen Zustand, der von einem intentionalen Zustand n-ter Stufe handelt, einen intentionalen Zustand n+1-ter Stufe. Alle intentionalen Zustände ab der zweiten Stufe wollen wir als meta-intentional bezeichnen.

Selbstkenntnis à la Descartes, wie sie oben charakterisiert wurde, ist meta-intentional; sie ist ein intentionaler Zustand mindestens zweiter Stufe. (Den gerade für Descartes interessanten Sonderfall selbstbezüglicher intentionaler Phänomene – jemand denkt zum Beispiel »Ich habe diesen Gedanken« als einen selbstbezüglichen Gedanken – lasse ich als unregelmäßige Merkwürdigkeit beiseite.¹⁴ Für solche Phänomene ist im obigen Schema keine Intentionalitätsstufe vorgesehen. Sie sind nicht auf der ersten Stufe anzusiedeln, denn sie handeln nicht von einem nicht-intentionalen Phänomen; sie sind auf keiner höheren Stufe als der ersten anzusiedeln, denn sie handeln nicht von einem intentionalen Phänomen niedrigerer Stufe.)

Der Repräsentationalist neigt dem Fehler zu, aus der Meta-Intentionalität auf Meta-Repräsentationalität zu schließen. In Analogie zu seinem charakteristischen Schluß

(1) X glaubt, daß es regnet

X hat in sich eine mentale Repräsentation, die zum Inhalt hat, daß es regnet

neigt er dazu, folgenden Schluß für gültig zu halten:

¹⁴ Siehe dazu Andreas Kemmerling, *Ideen des Ichs – Studien zu Descartes' Philosophie*, Frankfurt am Main 1996, Kapitel 2.

(2) X glaubt, daß er glaubt, daß es regnet

X hat in sich eine mentale Repräsentation, die zum Inhalt hat, daß er in sich eine mentale Repräsentation hat, die zum Inhalt hat, daß es regnet.

Nennen wir das dem Schluß (1) zugrunde liegende Folgerungsmuster das repräsentationalistische Schlußschema. Entsprechend liegt jedem Schluß vom Typ (2) das metarepräsentationalistische Schlußschema zugrunde; diese Bezeichnung weist darauf hin, daß in der Konklusion von (2) eine mentale Metarepräsentation postuliert wird, das heißt eine Repräsentation, die von einer Repräsentation als Repräsentation handelt. Es ist jedoch wichtig zu beachten, daß selbst dann, wenn (1) ein gültiger Schluß wäre, (2) immer noch ein ungültiger Schluß sein könnte. Der Metarepräsentationalismus folgt nicht aus dem Repräsentationalismus. Es ist völlig kohärent, (2) als ungültig abzulehnen, selbst wenn man irrigerweise (1) als gültig akzeptiert.¹⁵

Viele zeitgenössische Repräsentationalisten sind Metarepräsentationalisten im Hinblick auf meta-intentionale Phänomene. Genauer gesagt: Mir ist kein zeitgenössischer Repräsentationalist bekannt, der im Hinblick auf diese Phänomene nicht dem Metarepräsentationalismus anhing. Ein sehr verbreitetes metarepräsentationalistisches Modell der Selbstkenntnis ist das Scanning-Modell, das beispielsweise von W. Lycan vertreten wird.¹⁶ Es ist eine zeitgemäße Spielart der Lockeschen Lehre vom inneren Sinn. In der Variante von Lycan nimmt sie sich, kurz gefaßt, so aus: Subjekte, die Selbstkenntnis besitzen, haben innere Scanning-Vorrichtungen, mit denen sie ihre Geisteszustände ablesen können. Ein solcher Scanner ist »ein Aufmerksamkeitsmechanismus, der sich vermutlich auf repräsentationale Subsysteme [...] richten läßt«. ¹⁷ Introspektion besteht in solchen Ableseprozessen. Ein Subjekt hat »dieserlei inneren Zugang zu einigen seiner eigenen Geisteszustände erster Stufe und vielleicht zu einigen

15 Siehe dazu meine Arbeit »Selbstkenntnis ohne Metarepräsentationalität« (erscheint in: Julian Nida-Rümelin (Hg.), *Analysomen 3*, Berlin/New York 1999).

16 William G. Lycan, *Consciousness and Experience*, Cambridge, Mass. 1996, insbesondere Kapitel 2.

17 Ebd., S. 32.

höherer Stufe«. ¹⁸ Überzeugungen, die uns ja hier besonders interessieren, sind »innere repräsentationale Zustände des Nervensystems«. ¹⁹ Zu der von Lycan vertretenen Version des Repräsentationalismus gehört auch die These von der Existenz einer Sprache des Geistes. Eine mentale Repräsentation ist für ihn demnach ein Vorkommnis eines Ausdrucks der *lingua mentis* des betreffenden Subjekts, also ein konkretes Phänomen im Gehirn, das eine sprachartige Struktur aufweist (insbesondere Kompositionalitätscharakter besitzt). ²⁰ Die Funktionsweise eines inneren Scanners beschreibt er nun folgendermaßen:

Als unmittelbare Konsequenz der Tätigkeit eines inneren Scanners des Subjekts S erzeugt S ein Vorkommnis eines mentalen Worts für den Zustandstyp erster Stufe, der gerade abgelesen wird. Dieses Wort [...] wäre semantisch primitiv [...] es hätte gewiß nicht dieselbe Bedeutung wie irgendein primitiver oder zusammengesetzter Ausdruck der öffentlichen Sprache namens Deutsch. In diesem Sinne wäre es ein privater Name, [...] den nur sein Benutzer verwenden könnte, um das tatsächliche Bezugsobjekt zu benennen. ²¹

Das sind empirische Mutmaßungen, keine philosophischen Analysen. Zudem sind sie reich an Annahmen, die selbst im Rahmen des Repräsentationalismus überflüssig sind. Als da sind: erstens die Annahme, daß der Metarepräsentationalismus wahr ist; zweitens die Annahme, daß jeder wenigstens eine Sprache des Geistes besitzt (Lycan vermutet, daß jeder mehrere solcher Sprachen besitzt; eine darunter ist das Introspektorische ²²); und drittens die Annahme, bei Selbstkenntnis sei eine prinzipiell nicht in eine öffentliche Sprache unübersetzbare Privatterminologie im Spiel.

Noch schwerer wiegt, daß gerade diejenige Selbstkenntnis, die uns hier interessiert, in solch einem Ansatz gar keinen Platz hat. Innere Scanner können, wie Lycan ausdrücklich hervorhebt, ²³ falsche Ablesungen ergeben und auch das Ergebnis der Ablesung

¹⁸ Ebd., S. 48.

¹⁹ Ebd., S. 56.

²⁰ Eine sorgfältige und in ihrem Ergebnis sehr kritische Studie zur These von der Existenz einer Sprache des Geistes hat Katia Saporiti kürzlich vorgelegt: *Die Sprache des Geistes*, Berlin/New York 1997.

²¹ Lycan, *Consciousness and Experience*, a. a. O., S. 60; etwa dasselbe noch einmal auf S. 100 f.

²² Ebd., S. 101.

²³ Ebd., S. 33 f.

mit beeinflussen. Es ist demnach gar nicht mehr erkennbar, wie es unsere Testfall-Selbstkenntnis geben könnte, die ja nach Voraussetzung gegen Irrtum gefeit ist. Der Lycansche Denker könnte sich nicht dessen gewiß sein, daß er jetzt denkt, daß es regnet. Er müßte es als eine Möglichkeit anerkennen, daß es ihm nur so vorkommt, als denke er jetzt, daß es regnet; denn seines Erachtens könnte ja der Scanner, der ihm übermittelt, daß er jetzt denkt, daß es regnet, nicht richtig funktionieren. Ja, er müßte es wohl auch noch für möglich halten, daß es ihm nur so vorkommt, als komme es ihm jetzt so vor, daß er denkt, daß es regnet. Denn auch der Scanner, der ihm übermittelt, daß es ihm jetzt so vorkommt, als denke er, daß es regnet – auch dieser Scanner ist halt nur ein braves, kleines, fehleranfälliges Ablesemodul. Und immer so weiter. In Lycans naturalistischem Repräsentationalismus ist kein Platz für Selbstkenntnis.²⁴

Aus diesen Gründen ist eine Theorie der Selbstkenntnis wie die Lycans philosophisch unattraktiv. Zur Begriffsanalyse taugt sie nicht, weil in ihr überflüssig starke empirische Annahmen gemacht werden. Und noch einschneidender ist, daß sie letztlich sogar das Phänomen selbst leugnet: jene Selbstkenntnis (à la Descartes), die wir hier als Test des Naturalismus verwenden.

Dretskes Entwurf

Fred Dretske hat vor einigen Jahren eine entschieden weniger voraussetzungsreiche Skizze eines naturalistischen Ansatzes zu einer repräsentationalistischen Theorie der Selbstkenntnis vorgelegt. Darin beschränkt er sich auf Selbstkenntnis, die es mit Sinnesempfindungen zu tun hat, die wir momentan haben.²⁵ Dretskes Überlegungen zum Thema »Selbstkenntnis« sind deshalb in meinen Augen, philosophisch gesehen, erheblich interessanter als die Lycans, weil er (i) das Phänomen der Selbstkenntnis nicht bestreitet und weil er Selbstkenntnis im Rahmen eines re-

²⁴ Es sei denn, es würde ein Super-Scanner angenommen, der »garantiert« – das heißt wenigstens mit naturgesetzlicher Notwendigkeit – niemals falsch abliest. Doch solch einen Unsinn könnte ein aufrechter Naturalist wie Lycan natürlich nicht mitmachen.

²⁵ Fred Dretske, *Naturalizing the Mind*, Cambridge/London 1995, Kapitel 2.

präsentationalistischen Naturalismus zu analysieren versucht, der (ii) mit möglichst schwachen empirischen Annahmen auskommt. Zwar ist auch sein Ansatz metarepräsentationalistisch, aber nichts hängt daran. Der Metarepräsentationalismus ist bei Dretske nur ein Flüchtigkeitsfehler, der sich leicht ausmerzen ließe.²⁶

Zunächst möchte ich kurz die Theorieskizze Dretskes darstellen. Es geht, wie gesagt, um Selbstkenntnis, die sich auf »im selben Moment« stattfindende Sinnesempfindungen bezieht. (Dretske zieht den Terminus »introspektives Wissen« vor. Ich verwende diese beiden Ausdrücke im folgenden austauschbar.) Die reizvolle Idee hinter der Analyse von Dretske ist folgende: Introspektives Wissen über die eigenen momentanen Sinnesempfindungen wird nicht dadurch gewonnen, daß das Subjekt seine Aufmerksamkeit nach »innen« richtet, um dort Informationen zu gewinnen. Nach Dretskes Darstellung dieser Dinge werden keine inneren Scanner aktiviert und abgelesen; beim Erwerb dieses Wissens wird keine besondere »innere« Information verarbeitet; und schon gar nicht ist die Rede von einer besonderen »inneren« Sprache, für deren Ausdrücke es sinngleiche Ausdrücke einer öffentlichen Sprache gar nicht geben könnte. Vielmehr zieht das Subjekt, laut Dretske, aus nichts weiter als der Information, die ihm durch die Sinne geliefert wird, einen Schluß auf die einschlägige Beschaffenheit seiner eigenen momentanen Sinnesempfindung.

Ein Beispiel liefert der folgende Fall. Harvey schaut mein Hemd an und kommt auf diese Weise zu der Überzeugung, daß mein Hemd blau ist, und zu dem (Selbst-)Wissen, daß er es in diesem Moment als blau wahrnimmt. Nach Dretske gleicht dies dem folgenden Fall: Jemand richtet seine akustische Aufmerksamkeit in eine bestimmte Richtung und kommt auf diese Weise zu der Überzeugung, daß sein Hund bellt, und zu der weiteren Überzeugung, daß der Briefträger da ist. In solch einem Fall liegt das vor, was Dretske verschobene Wahrnehmung [displaced perception] nennt: Vermittels der Wahrnehmung des Hundegebells gelangt die betreffende Person zu der Überzeugung, daß der Briefträger da ist. Und angeblich ganz analog dazu: Vermittels

26 Siehe dazu meine Arbeit »How Self-Knowledge Can't be Naturalized«, in: *Philosophical Studies* 95 (1999), S. 311-328.

der Wahrnehmung meines Hemdes kommt Harvey zu der (selbstkenntniskonstitutiven) Überzeugung, daß er es als blau wahrnimmt.

Der Reiz dieser Dretskeschen Analogie liegt für mich in folgendem. In ihr werden einige Clichés zum Thema Selbstkenntnis (implizit) attackiert. Es bedarf keiner »inneren Wahrnehmung«, um zu Selbstkenntnis darüber zu gelangen, als was man die Umgebung wahrnimmt. Beim Erwerb von derlei sogenanntem »introspektivem« Wissen geht »der Blick«, das heißt die auf Informationserwerb gerichtete Aufmerksamkeit, nur nach außen. Die Hinwendung »nach innen« geschieht schlußfolgernd, im Denken. Selbstkenntnis dieser Art ist also weder »introspektiv« noch »direkt« oder »unmittelbar«.²⁷ Soweit sie eine Form der Wahrnehmung ist, ist sie eine (sei's auch verschobene) Außenwahrnehmung; und was den traditionellen Gegensatz zwischen »unmittelbar wahrgenommen« und »durch Schlußfolgerung gewonnen« angeht, gehört solcherlei Selbstkenntnis im Lichte der Analogie Dretskes deutlich in die zweite Rubrik. – Mit Dretskes Analogie lassen sich also drei (darunter zwei meines Erachtens falsche) repräsentationalistische Clichés über Selbstkenntnis vermeiden: das von der Meta-Repräsentationalität, das von der Gerichtetheit nach innen und das von der Unmittelbarkeit. Hingegen läßt sich in ihr jenes charakteristische Moment vielleicht bewahren, das in Lycans Ansatz verlorenging: jener »Cartesische« Aspekt einer unbezweifelbaren und dem Irrtum gar nicht erst ausgesetzten Einsicht in eine kontingente Wahrheit.

Doch wie reizvoll diese Analogie auch sein mag, es ist zu fragen, ob sie sich in dem Rahmen, den Dretske bietet, überhaupt ausführen läßt. Betrachten wir sie also noch einmal genauer. Da waren zwei Fälle, die Dretske zueinander in Analogie bringen möchte: der Fall mit der verschobenen Wahrnehmung der Ankunft des Briefträgers und der Fall mit dem introspektiven Erwerb von Wissen darüber, wie die Farbe eines Hemds wahrgenommen wird. Seine Analyse des ersten Falls ergibt folgendes:

27 Auch Dretske ergeht sich manchmal noch in den üblichen Floskeln über »nicht-inferentielles Bewußtsein« und »direktes Wissen«, das der Geist von dem habe, was sich in ihm gegenwärtig manifestiert (vgl. etwa: *Naturalizing the Mind*, a. a. O., S. 39). Aber seine eigenen Überlegungen weisen in eine ganz andere Richtung.

Schema I: Vershoben wahrnehmen

Beispiel

S nimmt mittels des Hundegebells genau dann verschoben wahr, daß der Briefträger da ist, wenn gilt:

- (1) S hört (»erlebt«) den Hund bellen;
- (2) das Hundegebell ist ein zuverlässiges Anzeichen dafür, daß der Briefträger da ist;
- (3) S glaubt, daß (2);
- (4) S gewinnt – dank (1)-(3) – die Überzeugung, daß der Briefträger da ist.

Versuchen wir nun, nach diesem Vorbild die intendierte Analogie zum Selbstkenntnisfall so getreulich wie möglich nachzuzeichnen.

Schema II: Introspektives Wissen erwerben

Beispiel

S erwirbt mittels –?– genau dann introspektives Wissen darüber, daß er das Hemd als blau sieht, wenn gilt:

- (1*) S sieht (»erlebt«) das Hemd als blau;
- (2*) –?– ist ein zuverlässiges Anzeichen dafür, daß S das Hemd als blau sieht;
- (3*) S glaubt, daß (2*);
- (4*) S gewinnt – dank (1*)-(3*) – die Überzeugung, daß er das Hemd als blau sieht.²⁸

²⁸ Dretske unterstellt, daß wenn S unter solchen Umständen glaubt, daß er das Hemd als blau sieht, er dann auch weiß, daß er das Hemd als blau sieht. (Ebd., S. 53 ff.) Deshalb spricht Dretske in diesem Zusammenhang von introspektivem Wissen. Ich halte diese Redeweise, wie eingangs erwähnt, für verfehlt. Doch auch schon (4*) selbst, worin es ja nur um die Überzeugung und noch nicht um Wissen geht, ist eher befremdlich als klarerweise wahr. Wer ein Hemd als blau sieht, würde zwar Fragen wie »Ist das Hemd blau?«, »Kommt dir das Hemd blau

Das ist Dretskes Analogie. Sie wirft gewiß mehr als eine Frage auf. Ich möchte mich auf folgende konzentrieren: Was genau entspricht dem Hundegebell, von dem im Beispiel für verschobenes Wahrnehmen die Rede ist, beim Erwerb introspektiven Wissens?

Gesucht ist nicht einfach ein sprachlicher Ausdruck, der an die Stelle von »-?-« im Beispiel für Schema II paßt. Ein Wort mag sich alsbald einstellen. Doch müßte es, um für unsere Zwecke dienlich zu sein, etwas bezeichnen, für das in Dretskes theoretischem Rahmen ein Platz vorgesehen ist. Dreierlei kommt in Frage:

- (a) das visuelle Erlebnis, von dem in (1*) die Rede ist;
- (b) der Inhalt dieses visuellen Erlebnisses;
- (c) die Tatsache, daß S dieses visuelle Erlebnis hat.

Ein weiterer Kandidat für die Ersetzung der Leerstelle in (2*) ist in Dretskes Ansatz nicht zu erkennen. Und die drei genannten Möglichkeiten schlagen fehl.

(a) ist nicht das Gesuchte, und zwar aus wenigstens zwei Gründen. Erstens ist jenes visuelle Erlebnis kein Analogon zum Hundegebell, von dem in (2) die Rede ist, obwohl die Einsetzung für »-?-« ja gerade dem Hundegebell aus dem Beispiel für Schema I entsprechen müßte. Dem visuellen Erlebnis beim Als-blau-Sehen in Schema II müßte in Schema I ein auditives Erlebnis beim Hören des Hundegebells entsprechen. Aber Hundegebell und Hörerlebnis sind, auch für Dretske, sehr verschiedene Dinge. Eine Analogie der Schemata I und II käme also mit (a) nicht

vor?« und auch »Hältst du das Hemd für blau?« wohl mit einem Ja beantworten – aber »Glaubst du, daß du das Hemd als blau siehst?« ist eine ganz andere und recht merkwürdige Frage. Fragte man mich das, während ich unter Normalbedingungen auf ein Hemd blicke, würde ich ohne weiteres weder mit einem Ja noch mit einem Nein antworten. In gewissem Sinne verstünde ich die Frage nicht, obwohl ich die strikt wörtliche Bedeutung des Fragesatzes natürlich verstehe. Doch normalerweise gilt: Wenn ein gewöhnlicher Sprecher des Deutschen nicht die Neigung hat, die Frage »Glaubst du, daß p?« ohne weiteres bejahend zu beantworten, dann hat er auch nicht die Überzeugung, daß p. – Kurz, ich halte schon (4*) für falsch und erst recht, daß die darin erwähnte Überzeugung Wissen konstituiert. Aber ich werde dieserlei Vorbehalte beiseite lassen.

zustande. – Zweitens wäre das Ergebnis einer (a)-entsprechenden Ausfüllung von (2*) auch abgesehen von jedweder intendierten Analogie inakzeptabel. Betrachten wir:

(2*a) Das Erlebnis, das S beim Das-Hemd-als-blau-Sehen hat, ist ein zuverlässiges Anzeichen dafür, daß S das Hemd als blau sieht.

Dies wäre auch für Dretske schwer zu schlucken. Denn der (angeblich) angezeigte Sachverhalt und das (angeblich) anzeigende Erlebnis stehen hier in einer begrifflich zu engen Beziehung. Anzeige in einem interessanten Sinn liegt hier nicht vor. Man vergleiche:

Der Zustand, in dem Harvey sich beim Einen-Schnupfen-Haben befindet, ist ein zuverlässiges Anzeichen dafür, daß Harvey einen Schnupfen hat.

In diesen Beispielen laufen Anzeigendes und Angezeigtes – wenn wir von metaphysischen Feinheiten absehen – auf dasselbe hinaus. Es sind Fälle von Selbstanzeige vom Schema »p zeigt an, daß p«. Ist jedoch Selbstanzeige überhaupt echte Anzeige? Diese Frage ist berechtigt. Wir müssen sie hier jedoch gar nicht beantworten. Es reicht die unkontroverse Beobachtung, daß (2) klarerweise von echter Anzeige handelt. Das Bellen des Hundes ist eines, die Ankunft des Briefträgers ist etwas klarerweise anderes. Solche Deutlichkeit des Unterschieds zwischen *Indicans* und *Indicatum* ginge mit (a) verloren – und damit vielleicht die Wahrheit von (2*), gewiß jedoch wiederum die intendierte Analogie.

(b) ist ebenfalls nicht das Gesuchte. Denn der Inhalt des visuellen Erlebnisses, das S beim Anblick des Hemdes hat, ist für Dretske das Blausein des Hemds. Aber das Blausein des Hemds ist nicht – wie eine entsprechende Ausfüllung der Leerstelle von (2*) es verlangen würde – ein zuverlässiges Anzeichen dafür, daß S das Hemd als blau sieht. Im Rahmen des Ansatzes von Dretske wäre so etwas in grotesker Weise falsch. Zuverlässige Anzeige ist eine Beziehung der objektiven Abhängigkeit: Das *Indicans* hängt vom *Indicatum* ab.²⁹ Aber das Blausein des Hemds hängt klarer-

²⁹ Siehe dazu: Fred Dretske, *Explaining Behavior*, Cambridge/London 1988, S. 54 ff.

weise nicht davon ab, ob und wie das einzelne Subjekt S dieses Hemd sieht.

(c) liefert nur Notationsvarianten zu (a). Anders gesagt, gegen

(2*c) Die Tatsache, daß S das Hemd als blau sieht, ist ein zuverlässiges Anzeichen dafür, daß S das Hemd als blau sieht

sind inhaltlich – also wiederum von allen metaphysischen Schlaumeiereien abgesehen – dieselben Einwände zu erheben wie gegen (2*a).³⁰

Schluß

Die beiden betrachteten Ansätze versagen vor jener kaum zu bestreitenden Form von Selbstkenntnis, die wir hier betrachtet haben. Sie sind, soweit ich sehe, die beiden derzeit philosophisch instruktivsten Spielarten des naturalistischen Repräsentationalismus, die sich diesem Thema überhaupt stellen. Man mag sich nun fragen, was am naturalistischen Repräsentationalismus dasjenige ist, das fehlt. Meine Antwort ist: Der Erzfehler liegt darin, intentionalen Zuständen im Rahmen des Repräsentationalismus beikommen zu wollen.

³⁰ Zu einer weniger hastigen Darstellung meiner Kritik an Dretskes Analogie von Selbstkenntnis und verschobener Wahrnehmung siehe meine in Anmerkung 26 erwähnte Arbeit.