

Andreas Kemmerling (München)

Theorie des Geistes ohne Vernunft –  
Überlegungen zu einem Versuch,  
den Rationalitätsbegriff als wertlos zu erweisen

Von Vernunft ist in der analytischen Philosophie unserer Tage wenig die Rede, zumindest nicht unter dieser Bezeichnung. Desto mehr von Rationalität. Der Begriff der Rationalität soll das auffangen, was an älteren Begriffen der Vernunft im Lichte der Standards wissenschaftlicher Forschung des 20. Jahrhunderts haltbar ist. Doch selbst über dieses Überbleibsel alter Vernunft wird neuerdings nicht immer sehr vorteilhaft geredet. Im folgenden werde ich darstellen, welche Kritik ein scharfsinniger Zeitgenosse kürzlich am Begriff der Rationalität geübt hat.<sup>1</sup> Seine Kritik läuft, um es vorwegzunehmen, darauf hinaus, daß er ein unhaltbarer, bestenfalls ein belangloser Begriff sei. In dieser Kritik nimmt er einen Standpunkt ein, dem eine Idee des 20. Jahrhunderts zugrunde liegt: die Idee, der menschliche Geist sei – zumindest was die Ausübung seiner kognitiven Funktionen angeht – einem Computer vergleichbar. Etwas genauer gesagt ist es die Idee, das Hirn lasse sich als die Hardware einer informationsverarbeitenden Maschine betrachten und Aufgabe der Kognitionswissenschaft sei es, in einer rein syntaktischen Theorie zu beschreiben, welche Symbolmanipulationsprozesse bei der sog. Kognition im Hirn stattfinden. Diese Auffassungen halte ich für grundlegend verfehlt. Doch Stichs Versuch einer szientistischen Demontage des Rationalitätsbegriffs ist widerlegbar, ohne daß dazu erst einmal die Unhaltbarkeit der Computermetapher gezeigt werden müßte. Der Auseinandersetzung mit der Argumentation von Stephen Stich möchte ich einige begriffliche Vorbetrachtungen über Rationalität, Intentionalität und eine besondere Art der Wertbezogenheit vorausschicken. Was ich dabei sage, richtet sich nicht gegen Stichs Argumentation, sondern soll an einige Merkmale des Begriffs erinnern, für dessen Abschaffung Stich plädiert.

Die alte Vernunft war ein Vermögen, Rationalität ist eine Qualität.

---

<sup>1</sup> Stephen P. Stich, *The Fragmentation of Reason*, Cambridge, Mass. 1990.

Über die Vernunft sprach man gerne so, als sei sie selbst so etwas wie ein Subjekt, das etwas tut; auch Kant spricht in vielerlei Weise von ihr wie von einer Person, die denkt, anschaut, empfindet, Gesetze gibt und sogar „dem . . . vernünftigen Wesen das Sollen vorschreibt“.<sup>2</sup> So spricht man heute nicht mehr über Rationalität. Rationalität ist eine Eigenschaft durch und durch, die allerdings von ganz unterschiedlichen Dingen ausgesagt wird: von Handlungen, von Entscheidungen, von Lebewesen, von Überzeugungen (Annahmen), von Wünschen (Zielen), von Strategien, von Bewertungen und von vielerlei anderm mehr.

Was also ist Rationalität? Was ist das Gemeinsame, das eine rationale Handlung, ein rationales Lebewesen und ein rationales Sonstwas sinnvoll miteinander verbindet? Darauf eine befriedigende Antwort geziemender Kürze zu geben, ist gewiß schwer, vielleicht sogar unmöglich — und mit Sicherheit nichts, was ich versuchen werde.

Womöglich ist die Familie der Rationalitätsprädikate nicht durch *eine* durchgängige inhaltliche Gemeinsamkeit semantisch verbunden. Die tiefen Begriffe neigen nun einmal dazu, schwer begreifliche — oder jedenfalls schwer zu explizierende — Ähnlichkeiten der unter sie fallenden Gegenstände einzufangen. Doch es gibt eine sich durchziehende Faser in dem Band, das die Verwendungen des Wortes „rational“ sei's auch nur locker miteinander verbindet, eine Faser, die besonderer Erwähnung wert ist. Es ist der Begriff des guten Schlusses.

## I. Rationalität und gutes Schließen

So also vom guten Schluß gleich am Anfang. Alle Rationalität hat, so denke ich, mit gutem Schließen zu tun. Rational ist eine *Entscheidung* oder eine *Handlung*, die sich als ein guter Schluß aus den Wünschen und Erwartungen des Handelnden betrachten läßt; rational ist eine *Überzeugung*, die sich als das Ergebnis eines guten Schlusses aus Wahrnehmungen und/oder andern rationalen Überzeugungen betrachten läßt; rational ist ein *Wunsch*, der sich als das Ergebnis eines guten Schlusses aus grundlegenden Wünschen und möglicherweise hinzutretenden rationalen Überzeugungen darstellen läßt. Und so weiter. Vielleicht noch eines: Rational ist eine *Kreatur* nur dann, wenn die Bewertungsdimension des guten Schließens sinnvoll auf sie anwendbar ist.

Es sind immer intentionale Zustände im Spiel, wenn von Rationalität

---

<sup>2</sup> Vgl. *Grundlegung zur Metaphysik der Sitten*, Akademieausgabe Bd. 4, Berlin 1903, S. 457 ff.

die Rede ist: Zustände von der Art wie, eine Überzeugung zu haben oder einen Wunsch zu haben. Auch das Schließen hat es wesentlich mit Intentionalität zu tun, es ist ja ein Übergang von einem intentionalen Zustand zu einem andern intentionalen Zustand. Rationalität kommt also nicht ohne Intentionalität daher; und übrigens auch, obwohl ich dazu nichts weiter sagen werde, umgekehrt Intentionalität nicht ohne Rationalität. Die beiden sitzen, begrifflich gesehen, im gleichen Boot. Gegen diese Auffassung mag es Vorbehalte geben. Sie ist jedoch innerhalb desjenigen Diskussionsrahmens unbestritten, um den es im folgenden gehen soll. Ich meine die philosophischen Überlegungen zum Begriff der Rationalität, wie sie während der letzten fünfzig Jahre innerhalb der sog. analytischen Philosophie angestellt wurden. Der Einfachheit halber spreche ich im folgenden von der *Verklammerungsthese*, um ein einziges Wort für die Behauptung zu haben, daß solch eine wechselseitige begriffliche Abhängigkeit von Rationalität und Intentionalität besteht. Um ihren Inhalt möglichst klar zu machen, möchte ich sie gleich noch einmal in Kontraposition formulieren: Wenn Rationalität eine Fiktion ist, dann ist auch Intentionalität eine; wenn Intentionalität eine ist, dann ist auch Rationalität eine Fiktion. — Diese Verklammerungsthese wird auch von Stich akzeptiert, dessen „Vernunftkritik“ ich mich im zweiten Teil dieser Arbeit zuwenden möchte; sie wird von ihm nicht nur akzeptiert, sie bildet den Ausgangspunkt seines Versuchs einer Demontage des Rationalitätsbegriffs. Doch zunächst möchte ich noch einige weitere vorbereitende begriffliche Betrachtungen anstellen, gegen die Stich — soweit ich sehe — wohl nichts einzuwenden hätte, sondern die er eher als Wasser auf seine Mühlen begrüßen könnte.

Ich hatte gesagt, eine Kreatur sei rational nur dann, wenn die Bewertungsdimension des guten Schließens sinnvoll auf sie anwendbar sei. Damit ist gemeint, daß vom Schließen nicht nur metaphorisch oder analogisierend die Rede sein darf (wie etwa vom Rechnen des Taschenrechners). Wenn man wissen will, ob die Bewertungsdimension des guten Schließens auf ein Subjekt überhaupt im eigentlichen Wortsinn anwendbar ist, dann ist es eine wichtige Frage, ob dieses Subjekt *tatsächlich* Überzeugungen und Wünsche hat. Denn eine Kreatur, die *sensu stricto* keine Überzeugungen hat, hält nichts für wahr (oder für falsch), nichts für wahrscheinlich (oder für unwahrscheinlich) und nichts für erstrebenswert (oder für vermeidenswert). Ihr fehlen mithin begriffliche Voraussetzungen, die fürs deduktive, induktive und praktische Schließen erforderlich sind. Und eine Kreatur, die keine Wünsche hat, versucht nichts. Nun ist Schließen aber der Versuch, für wertvoll gehaltene Eigenschaften der Prämissen auf die Konklusion hinüberzuretten. Zum wirklichen Schließen gehört also wirkliches Glauben und Wollen.

Der Begriff der Rationalität hat demnach, wie ich von nun an einfach

unterstelle, einen roten Faden.<sup>3</sup> Man sollte diesen Faden aber nicht für den Kern der Sache halten, schon gar nicht für einen harten. Denn was ist letztlich, wenn man es genauer nähme, gutes Schließen — oder gar *das* gute Schließen? Die deduktive Logik gibt es längst nicht mehr als einen echten Singular; wir haben heute den Pluralismus der deduktiven Logiksysteme. Bei der induktiven Logik liegen die Dinge gewissermaßen umgekehrt. Kaum jemand scheint inzwischen mehr daran zu glauben, daß es auch nur ein einziges System der induktiven Logik je geben wird. Bei der Entscheidungslogik und der Handlungslogik wiederum gibt es offenbar konkurrierende Ansätze und keine Aussicht auf den großen Konsens. Kurz, eine allgemeine und zugleich präzise und darüber hinaus auch noch des Beifalls der Klugen und Wohlmeinenden gewisse Theorie des guten Schlusses ist nicht nur nicht in Sicht; es wäre unvernünftig, auf dergleichen zu hoffen. „Deduktive Logik“, „induktive Logik“, „Präferenzlogik“, „Entscheidungslogik“ und was es sonst noch gibt, das sind Bezeichnungen für auf unterschiedliche Weise bearbeitete Problemfelder, nicht Bezeichnungen für Theorien, schon gar nicht für Theorien mit haltbarem Einzigkeitsanspruch.

Der demnach gebotene Theorie-Pessimismus, den übrigens Putnam gerne gegen den Computer-Funktionalismus ins Feld führt,<sup>4</sup> ist allerdings nichts, wodurch der Rationalitätsbegriff selbst gefährdet ist. Wenn gewisse Formen des guten Schließens gar nicht in einer systematischen (und letztlich formalen) Theorie erfaßbar sein sollten, dann besagt das weniger über die Grenzen der Vernunft als über die Grenzen dessen, was sich in einer formalen Theorie erfassen läßt. Und wenn gewisse andere Formen des guten Schließens gleichermaßen gut auf einander ausschließende Weisen formulierbar sein sollten, dann zeigt dies für sich genommen zunächst einmal nichts weiter, als daß auch die Theorien über Rationalität die übliche Theorien-Unterbestimmtheit aufweisen. Dadurch steht Rationalität jedoch in keiner Weise schlechter da als beliebige andere Gegenstände unserer Theoriebildung.

Von solcherlei Theorie-Pessimismus im Hinblick auf gutes Schließen — der ja nichts anderes ist als eine Absage an überzogene Hoffnungen auf eine allumfassende und garantiert konkurrenzlose Theorie der Rationalität — von solch vernünftigem Theorie-Pessimismus müssen Positionen unterschieden werden, die einen Skeptizismus im Hinblick auf das gute Schließen umfassen oder wenigstens einen entsprechenden Defätis-

---

3 Das gute Schließen ist nicht das einzige, worin man einen roten Faden der Rationalität erblicken kann. Der Begriff des guten Grunds z. B. hat nicht weniger Anrecht auf diesen Titel. (Es liegt auf der Hand, daß die beiden eng miteinander verwoben sind.)

4 Vgl. Hilary Putnam, *Representation and Reality*, Cambridge, Mass. 1988.

mus. Als skeptizistisch bezeichne ich solche Auffassungen, denen zufolge es so etwas wie gutes Schließen (also auch Rationalität) letztlich gar nicht gibt. Als defätistisch bezeichne ich Auffassungen, die besagen oder implizieren, daß der Begriff des guten Schließens (und mithin auch der der Rationalität) zwar nicht unbedingt leer, aber doch überflüssig, witzlos, ohne Wert sei. Davon später mehr.

Zunächst einmal stellt sich allerdings eine grundsätzliche Frage. Schließen *wir* denn eigentlich gut oder doch wenigstens leidlich gut? Schließen wir — jeder von uns einzeln betrachtet — gut genug, um uns wirklich die Feder der Rationalität an unsern Hut stecken zu dürfen? Die empirischen Befunde sind erschreckend. Einige Psychologen haben sich einen nachsenswerten Sport daraus gemacht, denkwürdige Denkfehler als denkbar weitverbreitet nachzuweisen. Sie haben mit geradezu maliziöser Ingeniosität Experimente entwickelt, die schlechtes Schließen aus ihren Versuchspersonen hervorkitzeln. Philosophen haben recht gut lachen bei derlei Aufgaben, weil es sich bei den schlechten Schlüssen z. B. um Verwechslungen wie die von „wenn . . . , dann . . .“ mit „genau dann . . . , wenn . . .“ und andere Dinge dieses Schlags handelt. Mit solchen Dingen mußten sich Philosophen seit dem ersten Semester traktieren lassen. Ein nicht speziell geschulter Mensch schneidet wahrscheinlich erst einmal schlecht ab. Machen wir — oder unter uns gesprochen: machen die andern — also zuhauf logische Fehler? Schließen viele Menschen, die kein formallogisches oder wahrscheinlichkeitstheoretisches Training hatten, so schlecht, daß man sie gar nicht mehr als gute Schließer betrachten kann? Verwirken sie, auf Grund der schlechten Testergebnisse, den Anspruch auf den Ehrentitel „rationales Lebewesen“? Natürlich nicht. So etwas möchte wohl keiner sagen.

Auch Stephen Stich nicht, der jedoch einen noch viel drastischeren philosophischen Schluß aus Überlegungen zieht, die ihren Ausgang bei dergleichen empirischen Beobachtungen nehmen. Er kommt zu dem „überraschenden Ergebnis, daß Wahrheit als eine Erkenntnistugend nicht ernst zu nehmen ist“.<sup>5</sup> Eines seiner Zwischenergebnisse, dem ich mich alsbald zuwenden möchte, besagt, daß es letztlich keinen wissenschaftlich greifbaren Unterschied zwischen gutem Schließen und schlechtem Schließen gibt. Wir machen im Vollzuge unserer alltäglichen Verrichtungen und Sprechakte zwar einen derartigen Unterschied, wissenschaftlich gesehen gibt es ihn aber nicht.

---

<sup>5</sup> Stich, a. a. O., S. 21.

## II. „Schließen“ ein Vortrefflichkeitsprädikat

Gegen die Idee, es gebe beim Schließen den Unterschied zwischen Gut und Schlecht nicht, liegt es natürlich nahe, den Einwand zu erheben, sie demontiere das Schließen selbst. Und ich denke, dieser Einwand trifft zu. Denn Schließen ist eines von den Dingen, denen ein Unterschied zwischen Gut und Schlecht in einem interessanten Sinn wesentlich ist. Es gehört zum Sinn des Wortes „schließen“, daß der Unterschied — oder ein Unterschied — zwischen Gut und Schlecht auf das Schließen anwendbar ist. Bevor ich dies ein wenig erläutere, möchte ich allerdings darauf hinweisen, daß auch damit noch kein vorweggenommener Einwand gegen die Position von Stich erhoben wird, die ich nachfolgend angreifen werde. Im Gegenteil, Stich könnte mir gelassen alles konzedieren, was ich in diesem Abschnitt sagen werde; er würde aber wohl — fast hätte ich ohne den Zusatz dieser Parenthese gesagt: — daraus den Schluß ziehen: Hinweg mit dem Begriff des Schließens.

In welchem interessanten Sinn ist dem Schließen ein Unterschied zwischen Gut und Schlecht wesentlich? Nun, es gibt Prädikate „F“, für die gilt: wenn es etwas gibt, das ein F ist, dann gibt es auch gute (oder jedenfalls leidlich gute) Fs. Nennen wir derartige Prädikate *Vortrefflichkeitsprädikate*.<sup>6</sup> Man beachte, daß es sich nicht bei jedem Fall, in dem ein Vortrefflichkeitsprädikat „F“ auf etwas zutrifft, um ein gutes oder einigermaßen gutes F handeln muß. Ganz im Gegenteil, es gehört gerade zum Sinn (oder zur gewöhnlichen Verwendung) solcher Prädikate, daß sie begrifflichen Platz für mißratene Exemplare lassen.

Vortrefflichkeitsprädikate sind mithin keine Wertprädikate, obwohl sie natürlich — wie fast jedes Prädikat — wertend verwendet werden können. Die semantische Pointe eines Vortrefflichkeitsprädikates „F“ ist nicht, daß aus „a ist ein F“ folgt „a ist gut“ oder „a ist ein gutes F“; und die pragmatische Pointe ist nicht, daß ernstgemeinte Äußerungen von Sätzen des Typs „a ist F“ immer auch Empfehlungen sind. Es gibt rein deskriptive Standardverwendungen von Sätzen mit Vortrefflichkeitsprädikaten.<sup>7</sup> Ihr innerer Bezug auf das Normative ist ein anderer. Damit

---

6 Auf die Existenz solcher Prädikate hat mich Paul Grice einmal in einem Gespräch aufmerksam gemacht; er sprach von einem „standard of excellence“, der mit ihnen verknüpft sei.

7 Ich unterstelle hier, daß ein Wertprädikat „W“ in einfachen Prädikationen des Typs „a ist ein W“ ceteris paribus empfehlenden Charakter hat. Wer aber sagt: „Diese Satzfolge dort ist ein Schluß“, der sagt damit nicht implizit: „Geh' hin und schließ' auch so!“ oder „Diese Zeichenfolge gibt einen guten Schluß wieder“; wer sagt: „Diese Zeichenfolge ist ein Satz“, der sagt damit nicht implizit: „Das ist ein gut gebauter Satz“ oder „Geh' hin und sprich auch so!“ — Auch Fehlschlüsse und schlechte Schlüsse sind Schlüsse. Auch ungrammatische und sprachlich unge-

„F“ auf einen Gegenstand  $x$  zutrifft, muß  $x$  zwar eine gewisse F-bezogene Qualität besitzen, die kein Nicht-F hat;  $x$  darf also (im Hinblick auf F) ein gewisses „absolutes“ Güte-Niveau nicht unterschreiten.<sup>8</sup> Aber  $x$  muß kein gutes F sein, um überhaupt ein F zu sein.

Eine genauere Explikation will ich hier nicht versuchen. Ein paar Charakteristika, die dabei zu berücksichtigen wären, seien allerdings kurz genannt: Für ein Vortrefflichkeitsprädikat „F“ ist es kennzeichnend, daß man es nicht meistern kann, ohne zugleich den Begriff von einem *guten F* zu meistern. Entsprechend wird ein Vortrefflichkeitsprädikat von verschiedenen Personen nur dann geteilt, wenn sie nicht nur weitgehend darüber einig sind, was ein F ist und was nicht, sondern einig auch darüber, was ein gutes F ist und was nicht. Ein Beispiel zur Erläuterung: Das Prädikat „Photographie“ ist kein Vortrefflichkeitsprädikat, denn zwei Personen können sich angesichts beliebiger ihnen vorgelegter Gegenstände völlig einig darüber sein, welcher Gegenstand jeweils eine Photographie ist, und zugleich herzlich uneinig darüber, welche darunter gute Photographien sind. Das Prädikat „Übersetzung aus dem Englischen ins Deutsche“ hingegen ist ein Vortrefflichkeitsprädikat. Denn wer nicht den Begriff der guten Übersetzung hat, hat gar keinen Begriff der Übersetzung, und wer weiß, was eine Übersetzung vom Englischen ins Deutsche ist, muß mit jedem andern, der dies weiß, im Hinblick auf viele Fälle darin übereinstimmen, ob ein tadelloses, ein zweifelhaftes oder ein mißratenes Exemplar einer solchen Übersetzung vorliegt.

Intersubjektive Übereinstimmung darf bei Vortrefflichkeitsprädikaten also nicht nur auf den Fall der deskriptiven Zuschreibung beschränkt sein. Bei Vortrefflichkeitsprädikaten ist die Frage, ob etwas ein gutes F ist, weniger eine Frage des persönlichen Geschmacks als eine der Begriffsbeherrschung: Der Begriff des Fs kommt sozusagen schon in seiner qualitätsbezogenen Auffächerung daher und läßt sich nur mitsamt derselben meistern. Verschwände in einer Sprachgemeinschaft der normative Konsens darüber, was gute Fs sind, ohne daß Streit darüber entstünde, was Fs sind, dann wäre „F“ in dieser Sprachgemeinschaft kein Vortrefflichkeitsprädikat mehr. — Den meines Erachtens entscheidenden Punkt möchte ich zum Abschluß dieser unzulänglichen Bemerkungen

---

schickte Sätze sind Sätze. „Schluß“ und „Satz“ sind zwar Vortrefflichkeitsprädikate, aber keine Wertprädikate.

<sup>8</sup> Man denke hier z. B. wiederum an das Vortrefflichkeitsprädikat „Satz“. Damit eine Wortfolge ein Satz ist, darf sie nicht schierer Wortsalat sein; sie muß ein gewisses Güte-Niveau haben. „Eigentliche daß du Wert As quoll“ ist eine Wortfolge, die diesen Standard nicht erreicht. „Dieser Satz, kein Verb“ ist ein Zweifelsfall. „Wir haben ein altes Auto gekauft und mit ihm fahren umher“ ist gut genug, um ein schlechter Satz zu sein.

noch einmal in einer ontologischen Redeweise so formulieren: die Existenz von Fs (insbesondere natürlich von schlechten Fs) ist parasitär gegenüber der Existenz guter (oder ziemlich gelungener) Fs.<sup>9</sup> In diesem Sinne sind Vortrefflichkeitsprädikate wesentlich *wertorientiert* (wenn auch, wie gesagt, nicht wertend).

Es liegt auf der Hand, daß der Begriff des Vortrefflichkeitsprädikats selbst Probleme in sich birgt: Was heißt denn schon „gut“? Kann es nicht alles Mögliche heißen, wenn ein F „gut“ genannt wird? Die eine Übersetzung ist gut, insofern sie in schönen Reimen abgefaßt ist und leidlich ähnlichen Sinn liefert; die andere Übersetzung ist nur gut, wenn sie den exakten Informationsgehalt des Originals wiedergibt. Man denke nur an so unterschiedliche Textsorten wie Zeitungsmeldungen, expressionistische Gedichte, Bedienungsanleitungen, juristische Verträge oder gar Witze. Welch unterschiedliche Ansprüche sind da jeweils an den Übersetzer zu richten. Wie soll es hier einen einheitlichen Sinn von „gute Übersetzung“ geben?<sup>10</sup> — Ein anderes und grundsätzlicheres Problem hat mit der Entscheidbarkeit der Frage zu tun, ob ein vorgegebenes Begriffswort überhaupt ein Vortrefflichkeitsprädikat ist oder nicht. Mein Freund z. B. ist der Auffassung, das Wort „Freund“ sei kein Vortrefflichkeitsprädikat; ich denke, es ist eines. Es ist nicht klar, wie eine objektive Entscheidung dieses Streits herbeigeführt werden kann.<sup>11</sup> Kurz gesagt, der Begriff des Vortrefflichkeitsprädikats ist in mehr als einer Hinsicht vage.

---

9 Daraus folgt, daß das zeitlich erste F (falls es das gab) ein leidlich gutes F gewesen sein mußte. Nota bene: ein leidlich gutes F zumindest *im Lichte der damaligen Standards*. Denn natürlich kann, was als ein gutes F gilt, von Kontext zu Kontext variieren. Diese Kontextabhängigkeit ist semantisch völlig harmlos, solange in jedem Kontext normativer Konsens herrscht und die Extension von „F“ immer dieselbe bleibt. Deshalb werde ich dies im folgenden einfach außer acht lassen.

10 Hier liegt m. E. in Wirklichkeit kein Problem, das seinen Grund im Begriff des Vortrefflichkeitsprädikats hat. Das angesprochene Problem hat seinen Grund vielmehr darin, daß der Übersetzungsbegriff selbst in einer gewissen Weise semantisch disparat ist. Es gibt einfach unterschiedliche Übersetzungsmetiers, von der Lyrik-Übersetzung bis zur Übersetzung von Bedienungsanleitungen; für diese Metiers sind jeweils unterschiedliche Vortrefflichkeitsstandards einschlägig. Deshalb liegt es nahe anzunehmen, daß es „den“ Begriff der Übersetzung gar nicht gibt, sondern eine Begriffsfamilie, deren Angehörige eigenständige (wiewohl verwandte) Vortrefflichkeitsbegriffe sind.

11 Auch dies kratzt nicht tief, nicht tiefer jedenfalls als bei vielen semantischen Prädikaten. Auch der Begriff der Synonymität, zum Beispiel, ist von dieser Schwierigkeit behaftet. Mein Freund glaubt, die Wörter „hager“ und „dürr“ seien synonym; ich glaube, daß zwischen ihnen ein Bedeutungsunterschied besteht. Es ist nicht klar, wie eine objektive Entscheidung dieses Streits herbeigeführt werden kann.



Aber er ist klar genug, um zu behaupten: Das Prädikat „Schluß“ ist eindeutig ein Vortrefflichkeitsprädikat. Etwas ist höchstens dann ein Schluß, wenn es ein leidlich guter Schluß ist *oder wenigstens einer Menge von Übergängen zuzuordnen ist, von denen ziemlich viele leidlich gute Schlüsse sind*. Auch diese hervorgehobene Klausel ist wiederum vage, aber sie weist auf ein weiteres interessantes Charakteristikum von Vortrefflichkeitsprädikaten hin. Sie erinnert nämlich daran, daß ein schlechter Schluß nur da anzutreffen ist, wo auch leidlich gute Schlüsse anzutreffen sind. Etwas, das unter andern Umständen gar kein Schluß gewesen wäre, ist manchmal ein (sei's auch noch so schlechter) Schluß. Anders gesagt: Ob ein bestimmter Übergang von einem kognitiven Zustand  $Z$  zu einem andern,  $Z^*$ , noch ein schlechter oder schon gar kein Schluß mehr ist, das mag davon abhängen, welche Übergänge das betreffende Subjekt in der Vergangenheit gemacht hat und in der Zukunft machen wird. Die Anwendbarkeit eines Vortrefflichkeitsprädikats kann also manchmal sehr kompliziert geregelt sein: Ob es bei einer bestimmten Gelegenheit anwendbar ist oder nicht, mag sich danach richten, ob das betreffende Ereignis (der betreffende Zustand usw.), von dem das Prädikat ausgesagt werden soll, sich in ein sehr komplexes und zeitlich weit ausgreifendes Muster einfügt. Und die Frage nach der Zugehörigkeit einer Sache zu einem Muster läßt, wie uns v. Savigny und Dennett kürzlich in sehr unterschiedlicher Weise vorgeführt haben, oft keine objektiv eindeutige Antwort zu.<sup>12</sup>

Vagheit kommt hier also auf wenigstens zwei Stufen zum Tragen, aber sie hat keine so desaströsen Auswirkungen, daß der Begriff des Vortrefflichkeitsprädikats von Anfang an zum Scheitern verurteilt wäre. Erstens mag es, wie gesagt, oftmals unentscheidbar sein, ob ein vorgegebenes Begriffswort ein Vortrefflichkeitsprädikat ist oder nicht. Dennoch gibt es jede Menge Begriffswörter, die klarerweise Vortrefflichkeitsprädikate sind. (Man denke nur an viele Artefakt-Wörter: „Humidor“, „Korkenzieher“, „Telephonapparat“, „Textverarbeitungsprogramm“.) Zweitens wird es selbst bei einem Begriffswort, das ganz klar ein Vortrefflichkeitsprädikat ist, oftmals unentscheidbar sein, ob es nun auf einen gegebenen Anwendungskandidaten zutrifft oder nicht. Doch das heißt nicht, daß es keine klaren Anwendungsfälle gibt. — Uns geht es hier vornehmlich ums Schließen, und im Hinblick darauf möchte ich (entsprechend den beiden genannten Vagheitsstufen) ohne weitere Diskussion zwei Eindeutigkeitsthesen vertreten: (1) Das Begriffswort „schließen“

<sup>12</sup> Siehe Eike v. Savigny, Wittgensteins „*Philosophische Untersuchungen*“, Band 1, Frankfurt a. M. 1988, S. 15–24; Daniel C. Dennett, „Real Patterns“, *Journal of Philosophy* 88 (1991), S. 27–51.

(oder „Schluß“) ist eindeutig ein Vortrefflichkeitsprädikat; (2) es gibt viele Fälle, in denen eindeutig ein guter bzw. eindeutig ein schlechter Schluß vorliegt.

Folglich — wegen (1) — ist jeder Angriff auf die Haltbarkeit des Unterschieds zwischen gutem und schlechtem Schließen unvermeidlich auch ein Angriff auf die Haltbarkeit des Begriffs des Schließens selbst. Und das wäre dann — im Lichte der Annahme, daß gutes Schließen der rote (oder wenigstens: ein roter) Faden von Rationalität ist — unvermeidlich auch eine Attacke auf den Begriff der Rationalität.

Wir haben bislang drei Aspekte des Begriffs der Rationalität hervorgehoben, die ich kurz zusammenfassen möchte. Erstens, gutes Schließen liefert einen roten Faden, der sich durch die vielfältige Anwendbarkeit des Rationalitätsbegriffs hindurchzieht. Zweitens, es gibt eine begriffliche Verklammerung von Rationalität und Intentionalität. Drittens, der Begriff des Schließens — der rote Faden der Rationalität — ist wesentlich wertorientiert.<sup>13</sup>

### III. Stichs defätistische Attacke auf den Rationalitätsbegriff

Stephen Stich greift den Begriff der Rationalität dadurch an, daß er den Unterschied zwischen gutem und schlechtem Schließen diskreditiert. Zunächst möchte ich zwei Thesen formulieren, die Stich so zwar nicht vertritt, die aber den Geist (sit venia verbo) seiner Auffassungen erkennen lassen. Ich zeige sozusagen Stichs Pistole erst einmal ohne den Schalldämpfer, den immer zu benutzen er klug genug ist.

---

<sup>13</sup> Es verlockt zu vermuten, daß alles Intentionale — oder wenigstens doch Intentionalität insgesamt — wesentlich wertorientiert ist. Alle grundlegenden intentionalen Prädikate scheinen jedenfalls Vortrefflichkeitsprädikate zu sein: „Überzeugung“, „Wunsch“, „Handlungsabsicht“ und, soweit ich sehe, auch viele weitere Prädikate, die man außerdem noch für intentionale Grundprädikate zu halten geneigt sein könnte. (Insbesondere natürlich auch das vieldeutige „Sinn“.) Die im intentionalen Bereich einschlägige Art der Vortrefflichkeit bezeichnen wir gewöhnlich mit den Wörtern „vernünftig“ und „sinnvoll“. Wenn nun aber das intentionale Vokabular tatsächlich durch und durch Vortrefflichkeitsvokabular ist, dann liegt darin vielleicht ein guter Grund für die Annahme, daß das hermeneutische Prinzip der wohlwollenden Interpretation nicht bloß irgendeine optimistische Grille ist. Dies Prinzip (oder diese Familie von Prinzipien) besagt, daß diejenige unter den nach Ausschöpfung aller empirischen Daten noch haltbaren Deutungen (eines Texts, einer Handlung) vorzuziehen sei, die ceteris paribus den Text oder die Handlung ihrem Sinn nach als am wenigsten unvollkommen darstellt. Daß ein derartiges Prinzip besonders gut zur auf Vernunft und Sinn orientierten Sphäre der Intentionalität paßt, ist zwar nicht zwingend, aber nahelegend.

(These 1) Es gibt keine begrifflich vorgegebene Einschränkung dafür, wie schlecht ein Schluß sein kann. Anders gesagt, Schlüsse können beliebig schlecht und dennoch Schlüsse sein.

(These 2) Es gibt mithin auch keine begrifflich vorgegebene Einschränkung dafür, wie vernünftig eine Person sein muß, um mit Hilfe intentionaler Begrifflichkeit beschreibbar zu sein. Anders gesagt, eine Person kann beliebig irrational und dennoch intentional beschreibbar sein.

Was Stich buchstäblich behauptet, klingt ganz anders. Er akzeptiert nämlich die Verklammerungsthese und räumt mithin ein, daß die intentionale Begrifflichkeit tatsächlich nur dann auf eine Person anwendbar ist, wenn diese Person rational ist. Und dazu, daß jemand rational ist, gehört auch für Stich, daß sie nicht allzu oft allzu schlecht schließt. Aber genau dies macht Rationalität (und die mit ihr unweigerlich einhergehende Intentionalität) in Stichts Augen wissenschaftlich gesehen uninteressant. Eine Wissenschaft vom menschlichen Geist braucht sich um Rationalität, gutes Schließen und alles, was dazugehört, gar nicht zu kümmern. Sie weiß es besser.

Denn daß ein Schluß nicht allzu schlecht sein darf, um überhaupt einer zu sein, oder daß eine Person nicht irrational sein kann, falls sie überhaupt etwas glaubt oder wünscht, dies ist Stich zufolge nur das Ergebnis einer wissenschaftlich inakzeptablen Art zu reden: nämlich mit Hilfe intentionaler Begrifflichkeit zu reden. Stich räumt ein, daß wir so reden. Das hat aber, so meint er, weiter nichts zu sagen.

Denn wissenschaftlich gesehen, im Lichte unseres besten Ansatzes zu einer Erforschung des Geistes, sei all dies irrelevant. Unser Vernunftbegriff ist zwar wesentlich im intentionalen Idiom verankert, doch dieses Idiom — unsere Begrifflichkeit des Glaubens, Wollens usw. — hat für Stich zwei miteinander zusammenhängende Defizite: Erstens gehöre es zu den Anwendungsvoraussetzungen dieser Begrifflichkeit, unnatürliche Grenzen zu ziehen, Grenzen, denen in der Welt nichts entspricht; zweitens gehöre es zu den Anwendungsvoraussetzungen dieser Begrifflichkeit, eine Perspektive einzunehmen, die in gewissem Sinn subjektiv ist und jedenfalls der objektiven Betrachtungsweise entgegensteht, wie sie in den Wissenschaften angestrebt wird.

Diese beiden Vorwürfe Stichts möchte ich kurz erläutern. Daß durch die intentionale Begrifflichkeit unnatürliche (eigentlich sogar: fiktive) Grenzen gezogen werden, zeigt sich für Stich sehr augenfällig daran, daß wir einer Person Rationalität absprechen müssen und mithin keine intentionalen Geisteszustände zuschreiben dürfen, wenn sie von den Standards für gutes Schließen (lange genug) drastisch abweicht. Stich stellt nun folgendes Gedankenexperiment an. Da sei eine Person, die exakt dieselben Überzeugungen hat wie Stich selbst, sich aber von ihm hinsichtlich der

Schlußfolgerungsprinzipien in einer winzigen Kleinigkeit unterscheidet. Wir sollen uns dabei – im Einklang mit der von Stich zugrunde gelegten computerorientierten Theorie der kognitiven Prozesse – Überzeugungen als neuronal realisierte Satzvorkommnisse einer Sprache des Geistes vorstellen; und zwar sind die Überzeugungen einer Person diejenigen Satzvorkommnisse, die eine bestimmte funktionale Rolle  $f$  im kognitiven Gesamtsystem der Person spielen; die Wünsche dieser Person sind dann entsprechend diejenigen Satzvorkommnisse, die eine bestimmte andere Rolle  $f^*$  spielen. Der Satzbegriff, der hier verwendet wird, ist rein syntaktisch; zwei syntaktisch gleiche Vorkommnisse sind Vorkommnisse desselben Satzes. (Der Kürze halber werde ich einfach von  $f$ -Sätzen sprechen, um damit Sätze zu bezeichnen, die die funktionale Rolle  $f$  spielen, von der wir annehmen wollen, sie sei typisch für Überzeugungen.)

Das Gedankenexperiment, das Stich anstellt, besagt also bis hierher im Lichte dieser neuen Betrachtungsweise etwa folgendes: Stich und eine weitere Person haben in ihrem kognitiven Gesamtsystem Vorkommnisse genau derselben  $f$ -Sätze; die Prinzipien, gemäß denen diese Sätze interagieren, sind jedoch bei Stich ein klein wenig anders als bei der zweiten Person. Und nun stellen wir uns vor, daß sich an diese beiden eine beliebig lange Reihe von Personen anschließt, wobei alle Nachbarn in dieser Reihe sich jeweils so zueinander verhalten wie Stich und sein Nachbar: ihre  $f$ -Sätze sind jeweils gleich, die Prinzipien der Satzinteraktion<sup>14</sup> unterscheiden sich jedoch. Je weiter wir in dieser Reihe fortschreiten, desto stärker weichen die Satzinteraktionsprinzipien der betreffenden Personen von denen ab, die bei Stich vorliegen. Von irgendeinem Punkt der Reihe an werden wir den nachfolgenden Personen Rationalität (und mit hin auch Überzeugungen) nicht mehr zusprechen können. Die Prinzipien sind zu wild, um noch als Schluß-Prinzipien gelten zu können.<sup>15</sup> Stich sieht jedoch keine wissenschaftliche Berechtigung für diese Einteilung. Er schreibt:<sup>16</sup>

An dieser imaginären Reihe ist folgendes interessant: Bei jedem Paar unmittelbarer Nachbarn handelt es sich um psychologisch gesehen *sehr* ähnliche Personen. Und vom Standpunkt eines rein formalen computer-orientierten Paradigmas aus betrachtet gibt es da keine

---

<sup>14</sup> Der Terminus „Prinzip der Satzinteraktion“ ist innerhalb dieser computerorientierten Theorie kognitiver Prozesse der Nachfolgerbegriff für „Schlußprinzip“ oder „Folgerungsregel“.

<sup>15</sup> Wenn jemand auf Grund irgendwelcher Prinzipien von dem Satz „Es regnet“ unmittelbar zu dem Satz „Es regnet nicht“ übergeht, dann nennen wir diese Prinzipien nicht Schluß-Prinzipien.

<sup>16</sup> Stich, a. a. O., S. 52.

interessanten oder bedeutsamen Einschnitte. Es gibt keine natürliche oder theoretisch wohlbegründete Aufteilung dieser Personen in zwei Klassen.

Sein Argument ist also dies: Wendet man intentionale Begrifflichkeit an, um die in dieser Reihe versammelten Personen zu beschreiben, dann muß man sie in wenigstens zwei Gruppen einteilen: nämlich in die Gruppe derer, auf die das intentionale Vokabular überhaupt anwendbar ist, und in die restlichen. Das ist jedoch unnatürlich und theoretisch haltlos. — Soweit Stichs Begründung seines Vorwurfs, durch das intentionale Idiom seien wir gehalten, unnatürliche Grenzen zu ziehen.

Der zweite Vorwurf hängt damit eng zusammen. Nach Stichs Auffassung ist es nämlich einfach so, daß der Begriff „Rationalität“ letztlich gar nichts anderes beinhaltet als „schließt so ähnlich wie ich bzw. so ähnlich wie wir“.<sup>17</sup> Wenn wir noch einmal auf die Personenreihe im gerade vorgestellten Gedankenexperiment zurückgreifen, dann müßten wir — so meint Stich — sagen: rational sind die Personen, die relativ nahe bei Stich stehen. Und das ist eine arbiträre Auszeichnung des betreffenden Ausschnitts der Reihe. Die Eigenschaft, sich in einer unendlichen Reihe relativ nahe bei einem durch die Theorie nicht ausgezeichneten Punkt zu befinden, ist nichts, was von theoretischem Interesse ist. Daß man seine eigene Umgebung eingrenzt und das andere ausgrenzt, ist nichts als Gruppen-Subjektivismus; theoretisch besehen genauso schlimm wie Individuen-Subjektivismus. Selbst wenn jeder von uns alle andern Menschen mit eingrenzte, wäre dies nichts weiter als Menschen-Subjektivismus oder Anthropozentrismus.

Wenn ich Stich hier recht verstehe, dann will er auf folgendes hinaus: In dieser Reihe lassen sich beliebige Grenzen ziehen; z. B. eine, die nur um Stich selbst herum verläuft (Subjekt), oder eine, die Stich und seinen Gen-Pool einbegreift und alles andere ausgrenzt (Familie), oder eine, die alle schwarzhäufigen Amerikaner einbegreift und den Rest ausgrenzt; oder eine, die noch irrelevanter ist. All diese Grenzziehungen sind wertlos, ja letztlich sogar hinderlich, wenn wir daran interessiert sind, was es mit Kognition, mit Geist, eigentlich auf sich hat. Und dies gilt auch für die Grenze, die der Begriff der Rationalität uns nahelegt. Der Begriff der Rationalität grenzt schon ein gewisses Stück der Reihe heraus, und er ist allein deshalb schon uninteressant: weil er — wie all die andern uninteressanten Begriffe — überhaupt etwas aus ihr herausgrenzt.

---

<sup>17</sup> Mit Vorgriff auf einen Terminus, den ich gegen Ende der Arbeit einführen werde, könnte man besser sagen: „hat ein ähnliches Folgerungsprofil wie ich bzw. wie wir“.

Wir können nun besser verstehen und genauer sagen, was Stich behaupten will. Ich hoffe, wir können nun sogar besser verstehen, was Stich — weil er klug ist — zwar nicht behauptet, wohl aber seinem Herzen folgend eigentlich gerne behaupten möchte. Mit den beiden oben genannten Thesen habe ich ihn als Rationalitätsskeptizisten ausgegeben; das ist so nicht korrekt. Denn er selbst stilisiert sich zum bekennenden Rationalitätsdefätisten, der sich des Unterschieds wohl bewußt ist, der ihn vom Skeptizismus trennt. Stich vertritt die folgenden Thesen:

(These A) Jede Abgrenzung innerhalb der kognitiven Prozesse zwischen solchen, die als „Schlüsse“ bezeichnet werden können, und solchen, die nicht so bezeichnet werden können, ist objektiv gesehen unhaltbar. Wir können solcherlei Abgrenzungen zwar machen — und wir tun das auch im Rahmen der Anwendung unseres intentionalen Vokabulars —, aber sie sind letztlich sachlich unfundiert.

(These B) Jede Abgrenzung innerhalb der *f*-Satz-Übergangsprinzipien ist objektiv gesehen arbiträr. Wir können solche arbiträren Grenzen zwar ziehen — und im Rahmen der Anwendung unseres intentionalen Vokabulars tun wir dies auch —, aber die gezogenen Grenzen sind dennoch letztlich sachlich unfundiert.

Diese beiden Thesen sind defätistisch, weil sich aus ihnen ergibt, daß der Begriff der Rationalität (und die gesamte intentionale Begrifflichkeit, in die Rationalität eingebettet ist) bei wissenschaftlichem Licht besehen überflüssig ist. Doch nicht nur das, das Reden von Rationalität (und allem, was dazugehört) ist auch sachlich verzerrend, denn die objektive Einheitlichkeit der kognitiven Prozesse wird dadurch terminologisch eskamotiert, daß manche als „rational“ zu bezeichnen sind und andere nicht.

Dieser Auffassung fehlt offenkundig nicht viel zum Rationalitätsskeptizismus, denn Rationalität gibt es ihr zufolge ja nicht — außer in einem gänzlich blassen Sinn. Damit es so etwas wie Rationalität in einem nennenswerten Sinn gibt, müßte der Unterscheidung zwischen „rational“ und „nicht rational“ ein echter Unterschied entsprechen; Stich meint jedoch, daß es in Wirklichkeit nur eine willkürliche Grenzziehung gibt — und das ist eben kein echter Unterschied. Dem Buchstaben nach ist Stichts Position nicht skeptizistisch; er behauptet nicht, das Wort „rational“ treffe auf nichts zu. Dem Geiste nach ist sie es sehr wohl, denn sie impliziert, daß der Unterscheidung zwischen dem Rationalen und seinem Komplement kein wirklicher Unterschied entspricht. So oder so, der Rationalitätsbegriff wäre ein grundlegend untauglicher Begriff für jede wissenschaftliche Theorie des Geistes.

#### IV. Eine unhaltbare Voraussetzung

Diese Argumentation Stichs ist auf vielerlei Weise angreifbar, und es mag besonders verlockend sein, den Versuch zu machen, immanente Widersprüche herauszuarbeiten oder wenigstens nachzuweisen, daß Stich sich in performative Selbstwidersprüche verstricken muß. Ganz gewiß gibt es gute Gründe, Stichs zugrunde liegende Auffassung davon, was eine wissenschaftliche Theorie des Geistes ist, für abwegig zu halten.<sup>18</sup> Auf all dies möchte ich hier nicht eingehen, sondern auf eine Voraussetzung der Argumentation hinweisen, die selbst im Rahmen der übrigen Auffassungen Stichs nicht haltbar zu sein scheint. Mein Einwand wird letztlich sein: der Begriff des kognitiven Zustands, durch den innerhalb einer kognitionswissenschaftlichen Theorie des Geistes die traditionelle intentionale Begrifflichkeit ersetzt werden soll, ist parasitär gegenüber eben dieser Begrifflichkeit. Anders gesagt: jede kognitionswissenschaftlich interessante Konzeption davon, was kognitive Zustände sind, ist darauf angewiesen, daß die Anwendbarkeit der intentionalen Begrifflichkeit ihr als Orientierung dient.

Ich möchte hier also — der Fortführung des Streits zuliebe — nicht von vornherein in Frage stellen, daß es sog. *kognitive Zustände* und *kognitive Prozesse* in dem Sinne gibt, den Stich mit diesen von ihm ins Spiel gebrachten Fachtermini verbindet. Er schreibt:<sup>19</sup>

Der Ausdruck „kognitiver Zustand“ soll sowohl „eigentliche“ Überzeugungen bezeichnen, als auch solche „überzeugungsartigen“ Zustände, die *genau wie* die eigentlichen Überzeugungen sind, *außer* daß es für sie keine intentionalen Beschreibungen gibt. Und der Ausdruck „kognitiver Prozeß“ soll sowohl „echte“ Schlüsse bezeichnen, als auch solche „schlußartigen“ Prozesse, in denen kognitive Zustände manipuliert werden, für die es keine intentionalen Beschreibungen gibt.

Was solche kognitive Zustände und Prozesse letztlich sind, sagt Stich detaillierter im fünften Kapitel:

---

<sup>18</sup> Eine vorzügliche Kritik dieser Art findet sich bei Katia Saporiti: *Die Sprache des Geistes — Repräsentationalistische und syntaktische Theorie des Geistes im Vergleich*, Teil II, Kap. 5 (Münchner Dissertation 1993). Saporitis Kritik richtet sich gegen Stichs ausführliche Verteidigung seiner Auffassung davon, was eine wissenschaftliche Theorie des Geistes ist, wie er sie in seinem früheren Buch *From Folk Psychology to Cognitive Science* (Cambridge, Mass. 1983) vorgestellt hat.

<sup>19</sup> Stich, a. a. O., S. 53 (die Hervorhebungen stammen von mir und sollen auf etwas hinweisen, das ich zwar nicht weiter thematisieren will, aber auch nicht unvermerkt lassen mag: hier liegt ein Hund begraben — und zwar auf einem Hundefriedhof).

... die Idee ist, daß Überzeugungen komplexe psychische Zustände sind, die man sich — wie Sätze — als aus einfacheren Bestandteilen zusammengesetzt vorstellen kann. Wir können nun Überzeugungsvorkommnisse [belief tokens] dadurch mit den wohlgeformten Formeln einer uninterpretierten formalen Sprache in Zusammenhang bringen, daß wir die Elemente, aus denen Überzeugungen gemacht sind, den Symbolen dieser Sprache zuordnen, und zwar so, daß dabei Wohlgeformtheit bewahrt wird. Ja, wir können Überzeugungsvorkommnisse als Inschriften der betreffenden wohlgeformten Formeln betrachten — Inschriften, die im Kode der Nervenzellen abgefaßt sind. Eine Überzeugung haben ist dann dies: ein Vorkommnis einer wohlgeformten Formel haben, die passend im Hirn gespeichert ist.<sup>20</sup>

All dies ist meines Erachtens allein schon deshalb völlig verfehlt, weil die Unterscheidung zwischen Vorkommnis und Typ auf Überzeugungen (und andere intentionale Zustände) gar nicht anwendbar ist. Für sog. "belief tokens" läßt die gewöhnliche Verwendung des Überzeugungsbegriffs gar keinen Platz. Die gegenteilige Auffassung, wie sie insbesondere im letzten Zitat formuliert wird, ist in meinen Augen ein ontologisches Mißverständnis. Dieses Mißverständnis ist inspiriert durch ein meines Erachtens völlig schiefes Bild vom menschlichen Geist bzw. von menschlicher Kognition, und zwar durch das Bild vom Geist als Computer (d. h. als einer symbolischen Datenverarbeitungsmaschine). Doch all dies soll uns hier, wie gesagt, nicht kümmern.

Der für unsern Zusammenhang entscheidende Fehler ist nämlich ein anderer. Denn selbst wenn wir annehmen könnten, jedes inhaltlich bestimmbare Überzeugtsein eines Menschen sei letztlich ein neurophysiologisch beschreibbares Muster der Interaktion von Nervenzellen im Hirn dieses Menschen, selbst dann ergäbe sich nichts, was den Rationalitätsdefätismus stützt. Zur Stützung bedürfte es nämlich folgender Annahme: es ließen sich derartige Neuro-Muster (wie ich sie der Kürze halber nennen möchte) ohne Rückgriff auf den Bereich des Intentionalen in kognitionswissenschaftlich interessanter Weise syntaktisch klassifizieren. Diese Annahme ist jedoch falsch.

Denn es gibt einfach zu viele Neuro-Muster. Grob geschätzt hundert Milliarden Nervenzellen enthält das menschliche Hirn, die meisten von ihnen haben zwischen tausend und zehntausend Synapsen. Suchte man in diesem (ohnehin unübersichtlichen) Geschehen nach einer syntaktischen Struktur, dann bestünde wenig Hoffnung darauf, daß man — bei hinreichend geduldigem Hinblicken — auf einen beständig sich verändernden

---

<sup>20</sup> Stich, a. a. O., S. 109.



den, uninterpretierten Text stößt. Das Problem wäre natürlich nicht, überhaupt einen Text zu finden; das Problem wäre, unter den unabsehbar vielen Kandidaten irgendwelche auszuzeichnen.

Die Auszeichnung sollte eine im Hinblick auf Kognition sein. Es gälte zunächst einmal, denjenigen Anteil des neuronalen Geschehens beiseite zu lassen, der mit Kognition nichts zu tun hat, um dann für den Rest eine Syntax zu entwerfen. Ein Vergleich mag hier hilfreich sein, obwohl die Größenordnung der Komplexität darin eine sehr viel bescheidenere ist. Denken wir an das Vorhaben, in der gesamten vokalischen Schallproduktion eines Menschen (inklusive Niesen, Husten, Keuchen, Brummen, Pfeifen, Räuspern, Kichern, Summen) den Sprachschall zu identifizieren und für dessen Erzeugnisse dann — ohne die Zuhilfenahme semantischer Hypothesen — eine syntaktische Beschreibung anzugeben. Dies würde jeder Linguist als lächerliches Ansinnen abweisen. Auf diesem Weg findet man zu keiner Syntax einer Sprache. Selbst mit der Unterstützung semantischer Hypothesen tun Linguisten sich sehr schwer, natürliche Sprachen syntaktisch zu beschreiben.

Und damit sind wir beim Kern der Schwierigkeit, vor der die Argumentation von Stich steht. Denn die eigentliche Schwierigkeit liegt nicht darin, daß das Hirn so unvorstellbar komplex ist. Sondern darin, daß keinem natürlich vorfindlichen System seine Syntax (sozusagen „werkseitig“) eingeschrieben ist. Eine eindeutig ersichtliche Syntax treffen wir höchstens da an, wo wir durch Festlegung eine eindeutige Syntax verliehen haben (und uns an sie erinnern). Das Geschehen in menschlichen Hirnen ist — selbst für uns, die wir solche Hirne haben — ein in der Natur vorgefundenes System, dem wir keine Syntax durch Festlegung verordnet haben. Wenn wir wissen wollen, wie die sog. Kognition vor sich geht, dann stehen wir vor dem menschlichen Hirn wie der Ochs vor der frisch gestrichenen Stalltür — und nicht wie der Logiker vor der frisch geputzten Wandtafel. Der Logiker an der Tafel kann einfach ad libitum stipulieren, was in seinem Kalkül eine wohlgeformte Formel ist. Was immer er durch Definition festlegt oder festgelegt hat, gilt. Sobald er sauber definiert hat, liegt denkbar eindeutig fest, was wohlgeformt ist und was nicht. In diesem Rahmen hat es dann guten Sinn, von wohlgeformten Formeln zu sprechen. Weil jemand eine Syntax eindeutig festgelegt hat.

Nichts dergleichen trifft auf das Hirn und seine Syntax zu. Was innerhalb des neuronalen Geschehens im Hirn als „Inschrift einer wohlgeformten Formel“ anzuerkennen wäre, das ist nicht nur nicht eindeutig festgelegt. Es ist gänzlich unbestimmt. Stellen wir uns zum Vergleich vor, die Weltmeere wären voll von Scrabble-Spielsteinen; die Meeresoberfläche wäre von diesen Buchstaben bedeckt, die von den Wellen durcheinandergewirbelt werden. Nun führte uns jemand auf einen hohen Berg, von dem aus sich das gesamte Meeresgeschehen überblicken ließe, und er forderte

uns auf: „Schreib mir die Syntax davon!“ Es mag uns dann trösten, daß wir immerhin noch vor einer viel leichteren Aufgabe stehen als der vor das Hirn gestellte Forscher, der „die“ Syntax des Hirngeschehens schreiben soll.

Angesichts des immensen Spielraums für mögliche syntaktische Beschreibungen bedarf es irgendwelcher Einschränkungen oder Vorgaben, um zu einer syntaktischen Beschreibung zu gelangen, die psychologisch oder kognitionswissenschaftlich *interessant* ist. All dies würde Stich wohl ohne weiteres einräumen. Der Dissens — und damit mein eigentlicher Einwand — beginnt mit folgendem Problem: Wie sind die Einschränkungen beschaffen, mit deren Hilfe sich aus allen möglichen syntaktischen Beschreibungen diejenigen ausgrenzen lassen, die für die Kognitionswissenschaft von Belang sind? (Diese Frage behandelt Stich nicht.)

Mein Einwand beruht darauf, daß an diesem Punkt nun kein Weg um Intentionalität und Rationalität herumführt. Er lautet: Jedes Bündel von Einschränkungen, mit dem sich die kognitionswissenschaftlich relevanten von den irrelevanten Syntaxen des Hirngeschehens absondern lassen, nimmt — sei's auch nur implizit — Bezug auf Rationalität und Intentionalität. Kurz gefaßt: Ohne Rationalität kein Kriterium für Kognition.

Slogans sind mißverständlich. Deshalb gleich zweierlei dazu, was dieser Einwand nicht besagt. Erstens besagt er nicht, daß es keine rein syntaktischen Theorien menschlicher Kognition geben kann. Zweitens besagt er auch nicht, daß ein brauchbares Kriterium für Kognition prinzipiell nicht in rein syntaktischer Begrifflichkeit formulierbar ist. Er besagt vielmehr, daß die Erscheinungsformen von Rationalität der sachliche Kern der Kognitionswissenschaft sind. Die kognitiven Leistungen und Prozesse rationaler Kreaturen haben unter all den Daten, die eine kognitionswissenschaftliche Theorie erklären muß, eine ausgezeichnete Stellung. Es sind die Daten, die eine Theorie erklären muß, um überhaupt eine kognitionswissenschaftliche (oder psychologische) zu sein. Der Begriff des guten Schließens (und alle weitere intentionale Begrifflichkeit) mag in der Ausformulierung einer syntaktischen Theorie der Kognition zwar beiseite gelassen werden. Aber die Muster des guten Schließens selbst (und weitere Muster, die für Rationalität charakteristisch sein mögen) müssen in solch einer Theorie erkennbar bleiben. Und dies in folgendem Sinn: Wenn rationale Kreaturen typischerweise den guten Schluß von A und B auf C machen, dann muß eine syntaktische Theorie der kognitiven Prozesse jener Kreaturen einen Übergang von zwei Formeln  $F_1$  und  $F_2$  auf eine dritte,  $F_3$ , vorsehen, und es muß möglich sein, der Abfolge dieser drei Formeln als „intentionale Interpretation“ den Schluß von A und B auf C zuzuordnen.

Daß Rationalität und Intentionalität den Kern der Sache ausmachen, von der die Kognitionswissenschaft handelt, räumt Stich sogar implizit

ein. Denn obgleich er in kognitionswissenschaftlichen Theorien kein intentionales Vokabular dulden möchte, besteht er doch darauf, daß die Fälle, in denen dieses Vokabular anwendbar ist — Fälle, in denen jemand etwas glaubt, etwas will, etwas anstrebt usw. —, in kognitionswissenschaftlichen Theorien erfaßt werden. Ich zitiere noch einmal jene verräterische Stelle:<sup>21</sup>

Der Ausdruck „kognitiver Zustand“ soll sowohl „eigentliche“ Überzeugungen bezeichnen, als auch solche „überzeugungsartigen“ Zustände, die genau wie die eigentlichen Überzeugungen sind, außer daß es für sie keine intentionalen Beschreibungen gibt.

Das, was Stich (wenn auch nur in Anführungszeichen) eine „eigentliche Überzeugung“ nennt, das soll also auf jeden Fall in seine neue Kategorie des sog. kognitiven Zustands fallen. Und das ist für diesen neuen Begriff des kognitiven Zustands wesentlich. Typische „eigentliche“ intentionale Zustände sind von vornherein, kraft Stipulation, auch sog. kognitive Zustände. Mehr wissen wir eigentlich nicht darüber, was ein kognitiver Zustand à la Stich eigentlich sein soll: nämlich irgend etwas, das entweder intentional beschreibbar ist oder aber aus dem Rahmen intentionaler Beschreibbarkeit herausfällt. Was den Rahmen intentionaler Beschreibbarkeit ausmacht, davon haben wir eine leidlich klare Vorstellung; wir nennen diese leidlich klare Vorstellung *Rationalität*. Bei dem, was aus diesem Rahmen herausfällt, haben wir nur sehr vage Ahnungen, worin dessen „Kognitivität“ bestehen möchte; hier schuldet uns Stich jede sachliche Bestimmung.

Drehen wir also den Spieß einmal herum (und entwickeln den Einwand noch einmal von vorne). Fragen wir: „Was, bitte schön, soll denn eigentlich ein kognitiver Zustand sein?“ Auf diese Frage bietet Stich sinngemäß nur folgende Antwort: „Ein kognitiver Zustand ist ein Zustand des Zentralnervensystems; er mag intentional unbeschreibbar sein, aber er ist das Vorkommnis einer syntaktisch wohlgeformten Formel, die sich aus einer kognitionswissenschaftlichen Theorie des Geistes ergibt. Die kognitionswissenschaftliche Theorie spezifiziert die dazugehörige Syntax.“ Das klingt fein. Nun sind aber im Hirngeschehen unbestimmt viele syntaktische Muster — unbestimmt viele Syntaxen — „realisiert“. Welche von ihnen sind kognitionswissenschaftlich relevant? Gewiß nicht alle. Denn wenn jede beliebige Syntax des neurophysiologischen Geschehens im Hirn einer Person als kognitionswissenschaftlich relevant akzeptiert würde, dann wäre eine Theorie des Geistes nichts weiter als eine beliebige Ordnung der neuronalen Aktivitäten in der Manier einer Syntax. So

---

<sup>21</sup> Stich, a. a. O., S. 53 (siehe dazu auch Anm. 13 oben).

etwas würde niemand behaupten wollen; denn unter den möglichen Neuro-Syntaxen des Hirns einer Person sind auch solche, deren wohlgeformte Formeln niemals eine intentionale Interpretation zulassen.<sup>22</sup>

Wollte man kognitionswissenschaftliche Adäquatheitskriterien für Neuro-Syntaxen formulieren, dann wäre das folgende sicherlich eines der grundlegenden. *Eine Neuro-Syntax, die kognitionswissenschaftlich adäquat ist, muß wenigstens das intentionale Profil normaler Subjekte einigermaßen getreu nachzeichnen können.* Was das heißen soll, sei kurz erläutert: Betrachten wir ein normales Subjekt S (wie insbesondere einen leidlich vernünftigen Menschen) zu einem bestimmten Zeitpunkt t. S hat eine Reihe von intentionalen Eigenschaften: Jeder einzelnen Überzeugung, die S zu t hat, entspricht solch eine intentionale Eigenschaft, desgleichen jedem Wunsch, und so weiter.<sup>23</sup> Beschränken wir uns auf diejenigen intentionalen Eigenschaften, die S zu t offenkundig hat (jedenfalls so offenkundig, daß es unter wohlinformierten, aufmerksamen Beobachtern keine ernsthafte Diskussion darüber gäbe, ob S diese Eigenschaften zu t hat oder nicht); und diese Eigenschaften wollen wir das *intentionale Profil* (von S zu t) nennen. — Zur Illustration ein Beispiel. Das intentionale Profil einer bestimmten normalen Person, die eines Montags morgens um halb acht vom Wecker aus dem Schlaf gerissen worden ist, mag u. a. folgende intentionalen Eigenschaften als markante Punkte haben: Glauben, daß es höchste Zeit für sie ist aufzustehen; Wünschen, daß sie wenigstens noch eine Stunde weiterschlafen könnte; Jedermann-Beneiden, der jetzt neben ihr liegt und bis zum sanften Erwachen weiterschlafen kann; Hoffen, daß gutes Wetter ist; Beabsichtigen, vor dem Duschen einen Kaffee zu trinken; und dergleichen mehr. Zu ihrem intentionalen Profil zu diesem Zeitpunkt gehören auch zahllose (auf Grund ihres beständigen Vorhandenseins) weniger markante intentionale Eigenschaften: Glauben, daß Harvey doof ist; Wünschen, daß ihr

---

22 Zur Erläuterung: Es sind gewiß syntaktische Beschreibungen  $S_1, \dots, S_n$  des neuronalen Geschehens in meinem Hirn (von der Geburt bis zum Tode) möglich, so daß diejenigen Hirnzustände, die Vorkommnisse wohlgeformter Formeln von  $S_i$  sind, nicht als Realisierungen meiner Überzeugungen betrachtet werden könnten. Denken wir nur an eine Neuro-Syntax, die das Hirn in 100 Regionen unterteilt und als wohlgeformte Formel nur eine Hirnaktivität zuläßt, bei der in all diesen Regionen Synapsen feuern. Und nun denken wir uns, daß ich zeit meines Lebens niemals eine Überzeugung zu einem Moment gewinne oder verliere, in dem in all jenen 100 Regionen Synapsen feuern. Keine einzige wohlgeformte Formel dieser Syntax hätte (in Anwendung auf mich) eine intentionale Interpretation.

23 Dahinter steckt nicht mehr als dies: Der Überzeugung, daß p, entspricht die Eigenschaft, p zu glauben; dem Wunsch, daß p, entspricht die Eigenschaft, p zu wünschen; und so weiter. — Das Reden von intentionalen Dingen (wie Überzeugungen und Wünschen) wird ersetzt durch Reden über intentionale Eigenschaften.

Mann mehr Geld verdiente; Hoffen, daß es in ihrer Umgebung sauber, zuverlässig, ordentlich und höflich zugeht.<sup>24</sup>

Jede Neuro-Syntax für das Hirn von S muß demnach folgende Auflage erfüllen, um kognitionswissenschaftlich interessant zu sein: Sie charakterisiert neuronale Aktivitäten im Hirn von S in einer solchen Weise als Formel-Vorkommnisse, daß sich zu jedem beliebigen Zeitpunkt den weit-aus meisten Eigenschaften des momentanen intentionalen Profils von S eine (möglicherweise komplexe) wohlgeformte Formel zuordnen läßt, die als die neurophysiologische Realisierung der betreffenden intentionalen Eigenschaft betrachtet werden kann. Entsprechend ließe sich ein Begriff des *Folgerungsprofils* (von S für den Zeitraum d) charakterisieren, und eine weitere Auflage an jede kognitionswissenschaftlich interessante Neuro-Syntax wäre dann, daß die syntaktischen „Satzinteraktionsprinzipien“ (wie Stich das nennt) es erlauben, das jeweilige Folgerungsprofil einigermaßen getreu nachzuzeichnen.

Dies wären Minimalanforderungen an eine kognitionswissenschaftlich adäquate Neuro-Syntax; weitere Forderungen müßten es dann damit zu tun haben, daß solch eine Syntax in fruchtbarer Weise auch auf Subjekte anwendbar ist, die aus dem Rahmen der intentionalen Beschreibung fallen. Dabei ist allerdings gar nicht klar, worin genau die „Fruchtbarkeit“ der Anwendung eigentlich bestehen soll.<sup>25</sup>

Warum ist die Nachzeichnung des intentionalen (und inferentiellen) Profils eine grundlegende Adäquatheitsvoraussetzung für kognitions-wissenschaftliche Neuro-Syntaxen? Die sachlich richtige Antwort lautet natürlich: „Weil Kognition, wie sie uns vertraut ist, durch und durch intentional ist.“ Wissen, Erinnern, Wiedererkennen, Planen und alle andern paradigmatischen Kognitionsphänomene sind intentionale Phänomene par excellence. Diese Antwort wird einen Intentionalitätsdefätisten natürlich nicht beeindrucken. Eine zweite Antwort, die zwar nicht den Kern der Sache trifft, könnte ihm immerhin Anlaß zum Nachdenken

---

24 *Prinzipiell* unverlierbare und insofern triviale intentionale Eigenschaften (wie Glauben, daß irgendetwas mit sich selbst identisch ist) sollte man nicht zum Profil rechnen; sie gehören zu den Voraussetzungen der Profilbeschreibung. — Zu den Schwierigkeiten bei der Abgrenzung dessen, was zum intentionalen Profil und was zu seinen Voraussetzungen gehört, vgl. John Searle, *Intentionality*, Kap. 5 (Cambridge 1983) und *The Rediscovery of Mind*, Kap. 8 (Cambridge, Mass. 1992).

25 Angenommen, es stünden uns zwei Neuro-Syntaxen zur Verfügung, die für normale, intentional beschreibbare Subjekte die Minimalanforderungen gleichermaßen gut erfüllten; was genau müßte solch eine Syntax  $\Sigma$  im Hinblick auf Subjekte leisten, die nicht intentional beschreibbar sind, um besser zu sein als ihre Rivalin  $\Sigma^*$ ? Das ist eine wichtige Frage für jeden, der solch einen syntaktischen Ansatz für verheißungsvoll hält. Eine Antwort darauf kenne ich nicht, ich wüßte nicht einmal, daß jemand die Frage stellt.

geben: „Die Nachzeichnung des intentionalen Profils ist deshalb unverzichtbar, weil (a) offenkundig nicht jede beliebige Neuro-Syntax kognitionswissenschaftlich relevant ist, und weil es (b) de facto schlicht und einfach keine andere plausible Idee gibt, wie sich die möglicherweise aufschlußreichen von den von vornherein uninteressanten Syntax-Spekulationen trennen lassen.“

Mein Einwand läuft also, zugespitzt formuliert, auf folgendes hinaus. Der Begriff des kognitiven Zustands ist offenkundig als eine Erweiterung und Verallgemeinerung des Begriffs des intentionalen Zustands gedacht; das geht aus dem gerade wiederholten Zitat deutlich genug hervor. Solch eine Erweiterung ist (zunächst einmal) unweigerlich darauf angewiesen, ihrer begrifflichen Basis weitgehend die Treue zu halten; andernfalls wäre da nichts, was sie erweiterte oder verallgemeinerte. „Die Treue halten“ muß in unserm Fall heißen: intentionales Profil nachzeichnen, so gut es eben geht. Begriffliche Treue ist nun aber unvereinbar mit Defätismus oder gar Skeptizismus gegenüber dem, was der Ausgangsbegriff beinhaltet.

Zwar fallen Späne, wo Erweiterungsbegriffe zur Anwendung gebracht werden; sie dürfen aber nicht alles hinweghobeln, woran mit den Ausgangsbegriffen gearbeitet wurde. Erweiterungsbegriffe müssen gegenüber ihren Ausgangsbegriffen begrifflich konservativ sein: möglichst vieles von ihnen bewahren.<sup>26</sup>

Für eine Position wie die von Stich stellt sich deshalb folgendes Dilemma. Einmal vorausgesetzt, daß intentionale Zustände sich von andern kognitiven Zuständen durch kein verfügbares syntaktisches Kriterium abgrenzen lassen.<sup>27</sup> Entweder ist dann der neue Begriff des kognitiven Zustands eine konservative Erweiterung des alten Begriffs vom intentionalen Zustand — in diesem Falle müssen den neuen Theorien die alten Muster, die sie möglichst weitgehend bewahren sollen, vorgezeichnet sein. Nach Voraussetzung kann eine Abgrenzung der zu bewahrenden Strukturen aber im Rahmen neuro-syntaktischer Terminologie allein nicht gelingen; es bedarf der intentionalen Begrifflichkeit, um vorzuzeichnen, was innerhalb der (sei's auch dann zu überschreitenden) Grenzen liegt.<sup>28</sup> Oder der Begriff des kognitiven Zustands ist ganz autonom gegen-

---

26 Das heißt aber nicht, daß der Ausgangsbegriff den „begrifflichen Kern“ des Erweiterungsbegriffs ausmachen oder eine anderweitig herausragende theoretische Rolle in der neuen Theorie spielen müßte. (Wenn jemand als einziger eine Platzreservierung hat, so kann er dennoch am Rande zu sitzen kommen.)

27 Genau dies gehört ja zum Ausgangspunkt von Stichts Argumentation. „Es gibt keine natürliche oder theoretisch wohlbegründete Aufteilung“, heißt es an einer Stelle, die ich oben zitiert habe.

28 Es verhält sich hier in mancher Hinsicht wie bei der Anfertigung einer Syntax-

über dem des intentionalen Zustands und geht das Risiko ein, intentionale Profile durchwegs nicht mehr erkennen zu lassen — dann ist er kognitionswissenschaftlich uninteressant. Denn seit Jahrtausenden kennen wir nur ein unumstrittenes Beispiel für „kognitive“ Zustände: die Überzeugungen, Wünsche, Erinnerungen und sonstigen intentionalen Zustände von leidlich rationalen Menschen.

---

Theorie für eine gewöhnliche Sprache wie das Deutsche. Der vortheoretische Ausgangsbegriff des ordentlichen Satzes ist für eine Syntaxtheorie zu eng, weil er Auflagen enthält, die den Sinn und vielleicht noch andere nicht-syntaktische Faktoren betreffen. Viele grammatisch korrekte Sätze werden keine ordentlichen Sätze im vortheoretischen Sinn mehr sein. Aber der vortheoretische Begriff des ordentlichen Satzes ist unverzichtbar, weil er der Syntax diejenige Satzmenge vorgibt, die in ihr möglichst weitgehend bewahrt bleiben soll. Andernfalls läuft sie Gefahr, keine Syntax des Deutschen zu sein. Denn nur eine Sprachbeschreibung, in der hinreichend viele Sätze dieser Ausgangsmenge als grammatisch korrekt anerkannt werden, ist akzeptabel als eine Syntax des Deutschen. Dies ist eine selbstverständliche Adäquatheitsbedingung für eine Syntax des Deutschen.