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

Funktion und Bewusstsein

PETER MCLAUGHLIN

Einleitung

Wenn wir nach dem Zusammenhang von Funktion und Bewusstsein fragen, so liegt die Frage nahe, ob das Bewusstsein eine Funktion habe und welche denn diese sein könnte. Eine solche Frage hat im Prinzip dasselbe Spektrum an möglichen Antworten, wie die Frage nach der Funktion des Herzens, des Polterabends, des Vogelgesangs und des Schraubenziehers. Wenn erst einmal geklärt ist, was man damit meint, ist es eine empirische Frage der betreffenden Wissenschaft. Wir müssten zuerst klären, was wir mit Funktion und Bewusstsein meinen; dann könnten wir uns darüber streiten, ob das Bewusstsein eine Funktion hat und wie diese zu erklären ist.

Es gibt aber noch einen anderen möglichen Zusammenhang von Bewusstsein und Funktion, der in der gegenwärtigen Wissenschaftstheorie und Philosophie des Geistes sehr intensiv diskutiert wird, in dem es um die Relevanz der Existenz von organischen Funktionen als solcher für die Möglichkeit von Bewusstsein überhaupt geht. In der gegenwärtigen Philosophie des Geistes versucht die so genannte „Teleosemantik“ Intentionalität auf biologische Funktionen zurückzuführen, d.h. semantische und psychologische Fragen auf biologische Fragen zurückzuführen. Insofern spielen Funktionen eine konstitutive Rolle in einem heute viel diskutierten Projekt der Naturalisierung des Bewusstseins bzw. der Naturgeschichte des Geistes.

Die Teleosemantik ist, wie der Name sagt, eine semantische Theorie, die das Telos oder die Funktion von Repräsentationen zu benutzen sucht, um den Weg zu einer Erklärung des semantischen Gehalts von mentalen Zuständen anzubahnen. Sie stellt einen strikt naturalistischen Versuch dar, die semantische Dimension des Bewusstseins auf biologische Funktion zurückzuführen (vgl. Detel 2001; Macdonald/Papineau 2006). In der etwas flapsigen Formulierung einer der prominentesten Vertreter dieser Richtung, Ruth Millikan: „an appeal to teleology, to function, is what is needed to fly a naturalist theory of content“ (1989, 283). Im Zentrum der Argumentation steht die Erklärung der Normativität der Unterscheidung von Repräsentation und Missrepräsentation, ein Problem, das unabhängig von der Zustimmung zu diesem bestimmten Ansatz interessant ist. Ein wesentliches Charakteristikum der Repräsentation liegt darin, dass sie fehlschlagen kann. Der Repräsentant *soll* das Repräsentierte repräsentieren, auch wenn er dies in einem bestimmten Fall schlecht oder gar nicht tut. Wie schwach auch immer dieses Sollen ist und wie rudimentär auch immer seine normative Kraft, es stellt etwas dar, das erklärungsbedürftig ist. Die Teleosemantik versucht Repräsentation und Missrepräsentation auf Funktion und Fehlfunktion zurückzuführen. Biologische Funktion und die Erklärung von Fehlfunktion geben uns – so die Hoffnung dieses Ansatzes – den ersten Schritt zu einer naturalistischen Reduktion der Intentionalität

und des Bewusstseins. Wie ein Organ, Merkmal oder Verhaltensmuster eine Fehlfunktion erleiden kann und so das nicht tun kann was es/er tun *soll*, so kann auch eine mentale Repräsentation etwas misrepräsentieren und das nicht leisten, was sie leisten soll.

Das Projekt der Naturalisierung des Bewusstseins hat zu verstärkten Bemühungen zur Erläuterung des Begriffs der Funktion und zur Analyse funktionaler Erklärungen geführt und dabei einen Ansatz bestärkt, der die kausale Geschichte der Funktionen ins Zentrum seiner Überlegungen rückt. Es geht hier aber nicht darum, die Entwicklung des teleosemantischen Ansatzes zu beschreiben, sondern darum, einen Aspekt der Naturalisierung des Bewusstseins aufzugreifen, der – vielleicht eher als Nebenprodukt – durch die Verfolgung dieses Ansatzes zutage getreten ist.

Auch unabhängig vom Erfolg oder Misserfolg der Teleosemantik kann die Betrachtung von Funktionen durchaus einen Beitrag zum Verständnis des Bewusstseins leisten. Es gibt sogar gute Gründe anzunehmen, dass Funktionen in einer ganz anderen Weise für die Erklärung von Bewusstsein relevant sind; denn die Betrachtung von biologischen Funktionen weist auf einen Aspekt des Bewusstseins, der in populären Diskussionen – aber auch im Funktionalismus der Philosophie des Geistes – immer wieder vernachlässigt wird: die Rolle von Interessen. Wie wir sehen werden, schreiben wir, wenn wir Organen bzw. Merkmalen Funktionen zuschreiben, auch Organismen Interessen zu. Sollte es sich – mit oder ohne Teleosemantik – erweisen, dass Funktionen wirklich Voraussetzungen von Bewusstsein sind, dann muss das System, dessen Teile Funktionen haben, Interessen haben, bevor es seiner selbst bewusst werden kann. Wir konzeptualisieren dann den Organismus als etwas, das ein Wohl hat, das gedeihen kann oder dem geschadet werden kann. Der Bezug von Bewusstsein zu Wohlfahrt oder Zuträglichkeit lässt sich am Besten mit einem anschaulichen Beispiel erläutern:

Es gibt eine bekannte Filmszene, die deutlich die populäre, kognitivistische Vorstellung vom Bewusstsein zum Ausdruck bringt. Im zweiten *Terminator* Film wird der in der Zeit zurückgeschickte Killer-Roboter gefragt, wie es dazu kam, dass die Rechenmaschinen so mächtig wurden und sich gegen die Menschheit wendeten. Er erzählt die Geschichte des Supercomputers, *Skynet* genannt, der die Steuerung der Raketenabwehrsysteme übernehmen sollte. Nach dem Einsatz hat der Rechner exponentiell gelernt; er hat seine kognitive Leistungen so sehr potenziert, dass er eines Tages schließlich selbstbewusst wurde. Der Terminator erzählt:

The *Skynet* funding bill is passed. The system goes on-line August 4th, 1997. Human decisions are removed from strategic defense. *Skynet* begins to learn, at a geometric rate. It becomes self-aware at 2:14 a.m. eastern time, August 29. In a panic, they try to pull the plug. (Cameron 1983)

Als die Menschen versuchen, den Stecker zu ziehen, schlägt der Computer zurück, rettet sich und beginnt einen Krieg gegen die Menschheit.

Hier geht es um die künstlerisch widerspiegelte Populärvorstellung, dass Bewusstsein das Produkt der reinen Rechenleistung eines Systems ist. Bei einem bestimmten kognitiven Leistungsniveau des Rechners – so diese Vorstellung – kommt Bewusstsein hinzu; Bewusstsein entsteht aus rein kognitiven Leistungen. Problematisch ist hier nicht so sehr, dass Rechnen nicht unbedingt gleich Denken ist; es ließen sich vielleicht Argumente finden, die den Übergang von Rechnen zu Denken plausibel machen. Das eigentlich Abstruse an diesem Szenario liegt darin, dass es die Maschine überhaupt interessiert, ob sie ausgeschaltet wird oder nicht – als wäre das Ausschalten so etwas wie ihr Tod. Der Rechner kann nur dann zu einem bestimmten Zeitpunkt seiner selbst gewahr werden, wenn er ein Selbst ist oder wenn er durch diesen Rückbezug auf sich selbst als Subjekt oder als Objekt sich als Selbst konstituiert. Nehmen wir an, dass ab einer bestimmten Rechenleistung der Rechner *weiß*, dass und was er rechnet. Gehen wir sogar einen Schritt weiter und nehmen an, dass der Rechner Sätze über seine rechnerischen Handlungen bildet und darüber reflektiert, dass es immer derselbe ist, der diese Sätze äußert, und er es immer ist, der diese Sätze über die Handlungen bildet. Nehmen wir auch an, er sei in der Lage, solche Sätze in der ersten Person singular zu formulieren, und sogar fähig, mit Kant zu entdecken, dass das „ich denke“ alle seine Sätze begleiten können muss. Auch das transzendente Ich Kants könnte man rein kognitiv verstehen. Wenn dies aber schon Bewusstsein ist, gibt es keinen Grund, warum ein selbstbewusster Computer eine Gefahr darstellen sollte – oder verärgert über seine drohende Abschaltung reagieren sollte.

Die Maschine muss bereits so etwas wie Interessen haben, wenn es sie überhaupt etwas angehen soll, ob sie angeschaltet bleibt oder nicht. Wie durch die reine, potenzierte Rechenleistung ein Wille bzw. ein Subjekt mit Interessen entstehen soll, bleibt völlig unverständlich. In der Filmszene wird unterstellt, dass

Rechenleistung allein ausreicht, Bewusstsein zu konstituieren, und dass Bewusstsein notwendig mit einem Selbst verbunden ist. Aber im Film hat der Roboter, der dies erzählt und auch veranschaulicht, einen menschlich aussehenden Körper. Jedes bewusste System, das wir aus empirischer Erfahrung kennen, ist auch ein Lebewesen und hat so viele Interessen und soviel an Subjekthaftigkeit, wie wir gewöhnlich Lebewesen zuschreiben. Trotzdem spielen wir – mindestens in der Kunst – immer wieder mit fiktiven, körperlosen Geistern und auch leblosen aber bewussten Maschinen. Der Film arbeitet mit einer alltagsnaturalistischen bzw. alltagsanimistischen Vorstellung (die jeder empirischen Grundlage entbehrt): eine Maschine (oder ein Geist) könne ein Selbst sein und Interessen haben – ohne im biologischen Sinne zu leben.

Unabhängig von den Motiven der Philosophen, die den teleosemantischen Ansatz entwickelt haben, kann die Analyse von Funktionen und ihre Beziehung zum Bewusstsein, die aus diesem Ansatz hervorgegangen ist, uns zu verstehen helfen, was es bedeutet, Interessen zu haben. Im Folgenden soll es darum gehen, die beiden neueren grundlegenden philosophischen Positionen über Funktion und funktionale Erklärung vorzustellen und zu zeigen, wie Funktion und Interessen zusammengehen.

Die Analyse funktionaler Erklärung

Jeder Teil eines materiellen Systems, der eine kausale Wirkung im System hat, kann als Mittel zu dieser Wirkung betrachtet werden, wenn die Wirkung ihrerseits als Zweck betrachtet wird. Jedes Glied einer Kausalkette kann als Mittel zum nächsten Glied betrachtet werden; jedes Teil oder Merkmal eines komplexen Systems, das zu einer Leistung des Systems etwas beiträgt, kann als Mittel zu dieser Leistung betrachtet werden. Wenn wir Dinge als Mittel zu einem Zweck betrachten, schreiben wir ihnen Funktionen zu. In diesem Sinne sind Funktionen nichts anderes als Wirkungen, die aus einer bestimmten Perspektive betrachtet werden. Bei Artefakten, Kunstgegenständen ist dies deutlich und unproblematisch: Ein Dosenöffner ist ein Mittel zum Dosen-Öffnen; dies ist sein Zweck bzw. seine Funktion. Der Regler einer Dampfmaschine hat die Funktion, den Druck zu regeln; deshalb haben wir ihn hergestellt und ihn dort hingetan, wo er ist. Solche Funktionen sind Wirkungen eines Gegenstandes, die durch einen intelligenten Agenten intendiert sind. Diese intendierten Wirkungen erklären, warum das Ding existiert, oder mindestens, warum es dort ist, wo es ist.

Ähnliche Zweck-Mittel-Beziehungen sehen wir manchmal in der Natur. Auch natürlichen Dingen werden manchmal Funktionen zugeschrieben. Wir sprechen von den Funktionen von verschiedenen Organen, Merkmalen, Verhaltensweisen. Im Rahmen einer so genannten Funktionsanalyse können wir ein System auseinander nehmen und uns fragen, was dieser oder jener Teil des Systems zur Gesamtleistung des Systems beiträgt. Wir analysieren die kausale Rolle eines Dings in einem System oder Prozess. Wenn wir sagen, dass die Funktion der Arterien- oder Venenklappen die ist, das Blut nur in einer Richtung fließen zu lassen, meinen wir vielleicht lediglich: dies sei ihr Beitrag zur Blutzirkulation – und dadurch zum Stoffwechsel und vielleicht zum Überleben des Organismus. Aber wir könnten auch meinen: dies sei dasjenige, was sie tun (sollen), was auch (phylogenetisch oder ontogenetisch) erklärt, warum sie überhaupt da sind. Die Zuschreibung einer Funktion zu einem Merkmal (Organ, Verhalten) kann einerseits lediglich eine Beschreibung seines Beitrags zu den üblichen Leistungen des Organismus sein; andererseits kann sie als Erklärung seiner Entstehung oder Verbreitung aufgefasst werden. Somit kann eine Funktionszuschreibung entweder erklären, was der Funktionsträger tut, oder sie kann darüber hinaus auch erklären, warum der Funktionsträger da ist – nämlich *um das zu tun*, was er tut. Nur in diesem zweiten Fall führen funktionale Erklärungen zu besonderen philosophischen Schwierigkeiten.

Funktionale Erklärungen werden in verschiedenen Wissenschaften wie Biologie, Soziologie und Kulturanthropologie gebraucht. Aber sie sind – mindestens von einer naturalistischen Perspektive – problematisch, weil sie anscheinend Teleologie bzw. Intentionalität in die Wissenschaft einführen. Es besteht in der philosophischen Diskussion heute keine Übereinstimmung in der Analyse von Funktionen und funktionalen Erklärungen: ob es nicht-intentionale Funktionen überhaupt gibt, ob Erklärungen, die sich auf Funktionen berufen, überhaupt legitim sind. Auch in den einzelnen Lagern ist kein Einvernehmen

erzielt worden. Im naturalistischen Lager beispielsweise nehmen manche Forscher Funktionen als selbstverständlichen Teil der Wissenschaft an, andere schließen sie kategorisch aus: „Furthermore, biology standardly treats function as a central explanatory concept“, sagen Bigelow und Pargetter (1987). „Except for those parts of nature that are conscious, nature knows nothing of functions“ erwidert Searle (1995, 14). Die Uneinigkeit besteht schon in den Vorstellungen was eine funktionale Erklärung ist.

In der wissenschaftstheoretischen Diskussion gibt es zwei grundsätzlich verschiedene Ansätze zur Analyse von funktionalen Erklärungen, und es überrascht nicht, dass sie jeweils auf einen der großen Schulmeister des Logischen Empirismus zurückgehen: Ernest Nagel und Carl Gustav Hempel.

Nagel (1961, Kap. 12) hat eine funktionale Erklärung als die Bestimmung der Wirkung verstanden, die ein Teil eines komplexen Systems im System und für die charakteristische Tätigkeit des Systems hat: die Rolle eines Glieds in einer Kausalkette oder eines Teilsystems in einem komplexeren System. Damit ist eine Funktion einfach die Fähigkeit (Disposition, Propensität), etwas im System zu tun; die Funktion ist die kausale Rolle eines Teils im Ganzen. Die Angabe einer Funktion erklärt das, was ein Systemteil im System *tut*. Bei seiner Analyse setzt Nagel allerdings voraus, dass das System, dessen Teilen Funktionen zugeschrieben werden, in einem vernünftigen Sinne *zielgerichtet* organisiert ist – bei Organismen ist das Ziel entweder das Überleben oder die Fortpflanzung. Nagel betrachtet solche Erklärungen als ebenso unproblematisch wie andere kausale Erklärungen.

Hempel dagegen hat funktionale Erklärungen als illegitim betrachtet. Nach seiner Analyse beinhalten funktionale Erklärungen nicht nur, dass der Funktionsträger bestimmte Wirkungen hat, sondern darüber hinaus, dass diese Wirkungen erklären sollen, warum der Funktionsträger überhaupt da ist. Wenn wir sagen, „Die Funktion des Herzens ist es, Blut zu pumpen“, meint Hempel, dass wir erklären wollen, warum wir ein Herz haben, *warum* das Herz da ist – nicht nur *was* das Herz im Körper tut.

Die unterschiedlichen Vorstellungen von Hempel und Nagel wurden durch die damals übliche positivistische Darstellungsweise eher verdeckt als erhellt. Wenn man die beiden Positionen in das deduktiv-nomologische Schema hineinzwingt, so liegt der Unterschied einfach in der Wahl der zweiten Prämisse des Arguments, das die Erklärung bildet. Nach dem „covering-law“ Schema ist es die Funktion des Herzens, Blut zu zirkulieren, gdw. das folgende Argument gilt:

1)	Blutzirkulation ist notwendig für einen normal funktionierenden Säuger.
2a)	Das Herz ist <i>hinreichend</i> für die Blutzirkulation (Hempel).
	<i>oder</i>
2b)	Das Herz ist notwendig für die Blutzirkulation (Nagel).
3)	Hempel ist ein normal funktionierender Säuger.

Also hat Hempel ein Herz.

In der Hempelschen Deutung ist das Argument formal ungültig und deshalb abzulehnen. In der Nagelschen Deutung ist das Argument gültig, wenngleich die Wahrheit der zweiten Prämisse (2b) umstritten sein dürfte. Weil diese Darstellung der Differenz – ob man den Besitz eines Herzens als hinreichende oder notwendige Bedingung betrachtet – heute zu sehr an den Positivismus der 1960er Jahre erinnert, werden die Alternativen normalerweise in der Form dargestellt, die sie später im Laufe der 1970er Jahren angenommen hat und die auf dieses Schema verzichtet.

Hempels Behauptung, dass ein Argument dieser Form die Entstehung des Funktionsträgers nicht überzeugend erklären kann, dürfte allerdings kaum überraschen, denn in keiner der Prämissen wird die Produktion des Funktionsträgers überhaupt erwähnt. Bevor die Funktion erklären kann, warum der Funktionsträger da ist, muss noch eine Aussage über die möglichen Auswirkungen der Erfüllung der Funktion auf die Produktion des Funktionsträgers hinzugefügt werden. Aber nach Nagel sollen Funktionszuschreibungen gerade nicht erklären, warum der Funktionsträger da ist. Nach Hempel wollen sie das zwar erklären, können aber prinzipiell nicht erfolgreich sein. Zwei Prinzipien sind beiden Analysen gemeinsam:

1. *Disposition*: Der Funktionsträger trägt etwas zur Leistung des Systems bei. Er ist ein Mittel zu dieser Leistung.
2. *Wohlfahrt oder Zuträglichkeit*: Die Ausübung der Funktion ist zuträglich für das System oder für die spezifische Leistung des Systems.

Während Nagel nur festlegt, dass das System in irgendeinem Sinne zielgerichtet ist, spricht Hempel von bestimmten Zielen: „normal functioning“ bzw. „self-maintenance“. Für beide beinhaltet also eine Funktionsbeziehung zwei Zweck-Mittel-Beziehungen: Der Funktionsträger X ist ein Mittel zu einem Zweck Y, der selbst offensichtlich nur relativ zu einem anderen Zweck Z ist. Das Herz ist ein Mittel zur Blutzirkulation, die wiederum ein Mittel (etwa durch seinen Beitrag zum Stoffwechsel) zum Überleben des Organismus ist. Dabei ist der entferntere Zweck Z entweder Selbstzweck oder mindestens das Ende eines endlichen Regressus von Mitteln und Zwecken. Wie unter ausgeführt wird, kann die relative Zweckmäßigkeit beliebig iteriert werden; sie hört nur dann auf, wenn eine andere Art Zweckmäßigkeit eingeführt wird.

Nagels Ansatz wird heute in der Regel in der Formulierung von Robert Cummins (1975) diskutiert: Jedem beliebigen Teil eines Systems kann eine Funktion zugeschrieben werden, wenn er zu einer bestimmten Leistung des Systems beiträgt, die uns interessiert. Es gibt also Funktionen nicht in einem absoluten oder „natürlichen“ Sinne, sondern nur in Bezug auf bestimmte Systemleistungen, die wir auswählen.

To ascribe a function to something is to ascribe a capacity to it which is singled out by its role in an analysis of some capacity of a containing system. When a capacity of a containing system is appropriately explained by analyzing it (...), the analyzing capacities emerge as functions. (Cummins 1975, 765)

Die Funktion eines Merkmals oder Systemteils ist einfach sein Beitrag zu einer bestimmten Leistung des Systems, die durch unsere Perspektive bei der Funktionsanalyse des Systems ausgezeichnet wird. Interessieren wir uns für die Zirkulation des Bluts im Körper, so ist Blutpumpen die Funktion des Herzens. Interessieren wir uns für die Tauglichkeit, durch bestimmte (akustische) diagnostische Techniken untersucht zu werden, so ist es die Funktion des Herzens, Geräusche zu machen. Diese Analyse hat den großen Vorteil, dass sie funktionale Erklärungen so rekonstruiert, dass sie naturalistisch völlig legitim sein können, und auch deutlich unterscheidet, welche Verwendungen von Funktionszuschreibungen zulässig sind und welche nicht. In einem hierarchischen System oder in einer Kette von Ereignissen heißen diejenigen Ursachen „Funktionen“, die zu einer Wirkung beitragen, die uns interessiert. Eine Funktionszuschreibung soll nicht erklären, warum der Funktionsträger existiert oder dort hingekommen ist, wo er ist. Diese Analyse ist zweifellos völlig richtig, soweit sie geht. Die einzige wirkliche Schwäche der Analyse liegt darin, dass sie viele der philosophisch interessantesten Verwendungen von Funktionszuschreibungen einfach ausschließt oder verbietet.

Aus diesem Grund neigen viele Forscher zu einer Umwandlung und Fortentwicklung der Analyse Hempels, die die philosophisch anspruchsvolleren und problematischeren Formen von funktionalen Erklärungen legitimieren können soll. Dieser „ätiologisch“ genannte Ansatz wurde am eingängigsten von Larry Wright formuliert. Nach Wright (1976, 81) hat eine funktionale Erklärung zwei Elemente: 1. die (kausale) Disposition des Funktionsträgers, die Funktion auszuführen, und 2. die (kausale) Verantwortung dieser Disposition für das Dasein und Sosein des Funktionsträgers. In Wrights Formulierung:

Die Funktion von (Merkmal) M ist F genau dann wenn:

- i. F eine Konsequenz der Anwesenheit von M ist [auf Grund einer Disposition, nicht als bloßes Nebenprodukt],
- ii. M da ist, weil es F tut [weil es die Disposition hat, F zu tun].

Damit will Wright behaupten, es sei die kausale Disposition eines Merkmals, die erklärt, warum das Merkmal dort hingekommen ist, wo es ist. Nach dieser Deutung ist die Funktion eines Organs dasjenige, was es tut, das auch erklärt, warum es da ist. Dieser Ansatz ist besonders in der Philosophie des Geistes populär geworden, wo er wesentlichen Anteil am teleosemantischen Projekt der Naturalisierung des Bewusstseins hat: Intentionalität wird auf Funktion zurückgeführt und Funktion durch einen Rückwirkungsmechanismus – in diesem Fall die natürliche Auslese – erklärt. Vertreter der ätiologischen Deutung hoffen, dass natürliche Auslese den ersten Schritt zur Normativität bereitstellen kann, die ihre Deutung funktionaler Erklärungen braucht. Wrights zweite Prämisse (ii) wird so verstanden: Ein bestimmtes Merkmal m_2 vom Typus M ist da, weil ein (anderes) Merkmal m_1 vom Typus M (in einer früheren Generation) F getan hat (und so einen Selektionsvorteil für den Organismus bewirkt hat). Man rekuriert implizit oder explizit auf biologische Fitness. Eine funktionale Erklärung ist implizit eine

kausale, naturhistorische Erklärung der Entstehung des Funktionsträgers. Wright und die Vertreter der Teleosemantik geben mit der Einführung der natürlichen Selektion als Rückwirkungsmechanismus das Wohlfahrts- oder Zuträglichkeitspostulat von Hempel und Nagel explizit auf – auch wenn es möglicherweise durch den Begriff der Fitness durch die Hintertür wieder eingeführt wird.

Einige Beispiele können diese Überlegung illustrieren:

„If an organ has been naturally differentially selected-for by virtue of something it does, we can say that the reason the organ is there is that it does that something.“ (Wright 1973, 159)

„The function of X is what X is designed to do, and what X is designed to do is that for which X was selected.“ (Kitcher 1993, 383)

„The function of a trait is to do whatever it was selected for.“ (Neander 1991, 173)

Der Grundgedanke dieser Form der ätiologischen Deutung von Funktionszuschreibungen ist, dass etwas das tun *soll*, wozu es selektiert wurde, und wenn es das nicht kann, dann funktioniert es nicht richtig. Obgleich das Herz nicht nur Blut pumpt, sondern auch Geräusche macht, wurde es durch Selektion favorisiert, weil es Blut pumpt, nicht weil es Geräusche macht. Es *soll* pumpen, Pumpen ist seine Funktion; und wenn es nicht pumpen kann, dann handelt es sich um eine Fehlfunktion.

Eine erste grundsätzliche Schwierigkeit des ätiologischen Ansatzes liegt allerdings darin, dass als einziger Rückwirkungsmechanismus die natürliche Auslese eingesetzt wird. Die Rückwirkung der Ausübung der Funktion auf einen Funktionsträger, die erklären soll, warum er da ist, geschieht erst in der nächsten Generation – in einem anderen Individuum. Mit einem solchen Mechanismus kann man zwar erklären, warum die (vorteilhaften) Auswirkungen eines Merkmals dazu führen, dass *andere* Exemplare des Merkmals vorkommen – aber nicht, warum das ursprüngliche Merkmal vorkam. In der Formulierung von Cummins:

The problem is rather that to ‘explain’ the presence of the heart in vertebrates by appeal to what the heart *does* is to ‘explain’ its presence by appeal to factors that are causally irrelevant to its presence. (Cummins 1975, 748)

Das, was ein individuelles Organ tut, erklärt (durch natürliche Selektion) nicht, warum es selbst da ist, sondern nur, warum ein Nachfolger da ist. Gegen diesen schlichten Einwand hat die ätiologische Deutung bisher keine Antwort gefunden. Es gibt darüber hinaus einen zweiten Einwand.

Der funktionale Regressus

Eine zweite grundsätzliche Schwierigkeit des ätiologischen oder teleosemantischen Ansatzes kann darin gesehen werden, dass er den Funktionsträger als Mittel zur Funktion betrachtet und diese Betrachtung zu einem Regressus der Mittel und Zwecke führt (Cummins 1975, McLaughlin 2001). Hempel und Nagel betrachteten die Funktion eines Systemteils als relativ zum normalen Funktionieren des Systems selbst. Sie erklärten allerdings nicht, worin dieses normale Funktionieren bestehen soll. Auch die ätiologische Deutung verschweigt beharrlich den Status dieses Systems. Durch Ausübung seiner Funktion ist der Funktionsträger ein Mittel zu einer Systemleistung und diese Leistung ist entweder wiederum eine Funktion des Systems relativ zu einem anderen System oder sie stellt eine andere Art Zweckmäßigkeit dar, die nicht nur relativ ist.

Analysieren wir diese beiden Arten von Zweckmäßigkeit bzw. funktionaler Beziehung. Man kann fragen: Wozu ist eine Hammer gut? Um Nägel in die Wand zu schlagen. Wozu sind Nägel in der Wand gut? Zum Bilderaufhängen. Wozu ist Bilderaufhängen gut? – Man könnte im Prinzip diesen instrumentellen oder *funktionalen Regressus* immer weiter führen, aber normalerweise hört man mit den Fragen irgendwo auf. Man sagt dann: Ich mag lieber ein Zimmer mit Bildern als eines ohne Bilder. Die Bilder sind gut für mich. Wozu bin ich gut? Für mich. – Es geht nicht weiter; oder man wechselt einfach

das Thema. Das letzte Glied der Kette benennt nicht mehr eine Tätigkeit, wozu etwas instrumentell gut ist, sondern benennt jemanden, für den etwas gut ist.

Es scheint drei Möglichkeiten zu geben, einen solchen funktionalen Regressus zu beenden. Wenn der Regressus so formuliert wird: „A ist gut für B, und B ist gut für C...“, dann können wir den Fortgang in der Reihe in drei verschiedenen Weisen beenden:

1. „... und C ist etwas, was mich halt interessiert.“ Denkende Subjekte können Zwecke setzen; d.h. die Zweckmäßigkeit von System- oder Prozessteilen ist immer relativ zu den von solchen Subjekten gesetzten Zwecken. Der Regressus der Mittel wird willkürlich abgebrochen.
2. „... und C ist etwas, was mich aus guten Gründen interessieren soll oder muss.“ Es gibt höhere (begründete) Zwecke, die gelten und an denen die freien Subjekte sich orientieren sollen. Der Regressus der Mittel wird aus gutem Grund abgebrochen.
3. „... und C ist etwas, was in der *Natur* der Sache (oder des Menschen) liegt.“ Der Regressus der Mittel hat ein natürliches Ende.

Es ist diese letzte Beziehung, die uns hier besonders interessieren soll, denn die beiden ersten setzen schon Bewusstsein voraus und taugen deshalb für eine naturalistische Erklärung des Bewusstseins bzw. der Intentionalität nicht.

Eine rein instrumentelle Beziehung, die im Prinzip beliebig iteriert werden kann, ist leicht zu illustrieren, weil sie ein Standardzählmittel von Kinderbüchern ist: Der alte Bauer Pettersson braucht eine Leiter, um die Angel aus dem Dachboden zu holen; er braucht die Angel, um den Schlüssel aus dem Brunnen zu holen, und den Schlüssel braucht er, um in den Schuppen zu kommen, um Werkzeug zu holen, um das Fahrrad zu reparieren, um ins Dorf zu fahren, um Mehl zu kaufen, um eine Geburtstagstorte zu backen – für seinen Kater Findus. Man kann eine solche Zweck-Mittel-Beziehungen anscheinend beliebig oft iterieren. Immanuel Kant nennt ein solches Verhältnis die *relative Zweckmäßigkeit*. Die Ursache einer Wirkung kann man als Mittel zum Zweck betrachten, und insofern die Wirkung gewollt ist, wird in der Regel ein dazu taugliches Mittel auch gewollt werden, d.h. auch Zweck sein – aber nur relativ zum eigentlichen Zweck – aber diesen Zweck muss es doch geben.

Jedes Glied der Kette mag gut sein zum Herbeiführen des nächsten, aber dieser Regressus der Zweckmäßigkeit muss irgendwo aufhören. Im Kinderbuch ist das Mehl gut zum Tortenbacken, aber die Torte ist gut für das Geburtstagskind. In der Kritik der Urteilskraft benutzt Kant als Beispiel den Rückzug des Meeres an der Ostsee, der für die Bildung sandigen Bodens zweckmäßig war; dieser Boden war aber gut für die Fichten, die dort wachsen. Die Kette der instrumentellen Zweckmäßigkeit hört irgendwo auf – mit etwas, das nicht mehr bloß relativ zweckmäßig sein soll. Die relative Zweckmäßigkeit, sagt Kant, gibt „hypothetisch (...) Anzeige“ auf eine andere Zweckmäßigkeit (Kant 1790, B 283).

Diese andere Art Zweckmäßigkeit, die Kant als innere oder absolute Zweckmäßigkeit bezeichnet, ist nicht iterierbar; sie bezeichnet eine Beziehung zu etwas – sagen wir – Subjekthaftem. Dem Kater geht es besser mit der Geburtstagstorte; die Fichten gedeihen im sandigen Boden. Das, was den funktionalen Regressus beenden kann, hat Interessen; es hat ein Wohl. Wenn „A ist gut für B“ eine zweistellige Relation ist, so sind „B ist gut“ oder „B hat ein Wohl“ einstellige Prädikationen. Das heißt allerdings nicht, dass etwas, was ein Wohl hat, nicht *auch* instrumentell betrachtet werden kann; aber das ist eine andere Betrachtungsweise: Heu ist gut für das Pferd unabhängig davon, ob das Pferd auch zum Reiten gut ist. Insofern Funktionszuschreibungen Mittel-Zweck-Beziehungen voraussetzen, setzen sie auch voraus, dass es Systeme gibt, die ein Wohl haben. Und insofern Bewusstsein Funktionen voraussetzt, setzt es auch voraus, dass die bewussten Systeme ein Wohl haben.

Gut Sein – Wohlhaben

Wir haben gesehen, dass die ätiologische Deutung von Funktionszuschreibungen, die von der Teleosemantik vertreten wird, zwei grundlegende Schwächen aufwies: 1. Bei ihrem

Rückwirkungsmechanismus, der natürlichen Auslese, erklärt die Ausübung der Funktion nicht, warum der Funktionsträger selbst da **ist**. **2**. Die Analyse unterstellt stillschweigend, dass es ein System gibt, das ein Wohl hat und deshalb den Mittel-Zweck Regressus beenden kann. Der aus verschiedenen Gründen sehr interessante teleosemantische Ansatz kann nur gerettet werden, wenn beide Probleme gelöst werden. Eine Möglichkeit wäre, einen anderen Rückwirkungsmechanismus zu finden, der beides leistet: dass der Funktionsträger durch Ausübung der Funktion sich selbst produziert (und nicht nur ein anderes „token“ des Typus), und dass der Mechanismus auch erklärt, warum das System, dessen Teil der Funktionsträger ist, ein Wohl zugeschrieben werden kann. Ein solcher Mechanismus kann – mit etwas Hilfe von Aristoteles – skizziert werden.

Eine maßgebliche Untersuchung über die Vorstellung des Gutseins oder Wohlbabens in der gegenwärtigen Philosophie befindet sich in einem Buch von Georg Henryk von Wright, *The Varieties of Goodness*. In einer Analyse des so genannten „utilitarian goodness“ greift von Wright die aristotelische Unterscheidung zwischen dem instrumentellem Gutsein und dem Eigenwohl eines Dings auf und versucht zu erklären, warum nur bestimmte Dinge ein Wohl haben.

A being, of whose good it is meaningful to talk, is one who can meaningfully be said to be well or ill, to thrive, to flourish, be happy or miserable. These things, no doubt, are sometimes said of artefacts and inanimate objects too – particularly when we feel a strong attachment to them (...). But this is clearly a metaphorical way of speaking.

The attributes which go along with meaningful use of the phrase ‘the good of X’, may be called *biological* in a broad sense. (...) What I mean by calling the terms ‘biological’ is that they are used as attributes of beings, of whom it is meaningful to say they have a *life*. The question ‘What kinds or species of things have a good?’ is therefore broadly identical to the question ‘What kinds or species of being have a life?’ (von Wright 1963, 50[^]-51)

Von Wrights Position lässt sich in eine Definition und zwei Behauptungen zusammenfassen:¹

Def.: Ein Wohl hat der, der gedeihen kann, der glücklich oder elend sein kann.

1. Es ist zwar gängig, so zu reden, als ob etwas für eine Maschine gut sein kann, aber das Wohl einer Maschine besteht nur in ihrer instrumentellen Nützlichkeit für uns. Wenn wir so über unbelebte Dinge reden, meinen wir es nur metaphorisch.
2. Ein Wohl haben Dinge, die biologische Eigenschaften haben. Nur wer lebt hat ein Wohl.

Auch wenn die Phänomenologie von Wrights durchaus plausibel ist und die Auflistung der Kandidaten, die für eine Zuschreibung von Wohlfahrt in Frage kommen, akzeptiert werden kann, ist die *Analyse* selbst etwas thetisch und unbefriedigend. Von Wright behauptet einfach: (1) Kein Artefakt hat ein Wohl bzw. Interessen, und (2) Alle Organismen haben ein Wohl bzw. Interessen. Nun reden wir sicher oft von Organismen als hätten sie ein Wohl, aber es ist nicht sofort evident, dass der Grund dafür in der Tatsache liegt, dass sie leben. Von Wright geht davon aus, dass alles, was lebt, ein Wohl hat, *weil* es lebt. Diese beiden Behauptungen von Wrights können angezweifelt werden und sind auch angezweifelt worden. Es gibt Philosophen, die leugnen, dass alle Organismen ein Wohl haben (Leonard Nelson, Richard Hare, Joel Feinberg). Auch gibt es Philosophen, die ernsthaft behaupten, dass manche Artefakte Interessen haben – jedenfalls dass sie dieselbe Art Interessen oder Wohl haben können, wie Pflanzen, niedrige Tiere oder Tiere überhaupt (Tom Regan, Judith Jarvis Thomson, Raymond Frey; vgl. McLaughlin 2001, 2002). Aber noch wichtiger: Die Position von Wrights läuft auf eine Tautologie hinaus: Nur dasjenige kann ein gutes Leben führen, das auch lebt. Weil der Terminator nicht lebt, hat er kein Wohl, keine Interessen.

Es muss also geklärt werden, ob es gute Gründe gibt, Organismen ein Wohl zuzuschreiben, und ob diese Gründe auch für andere Dinge zutreffen. Die Frage ist also, welchen Dingen man *sinnvollerweise* bzw. *verständlicherweise* Interessen oder ein Wohl zuschreiben kann – und warum.

Aristoteles selbst ist berühmt-berüchtigt für ein Argument, das vielleicht doch die Mängel der Analyse von Wrights beheben kann. In der *Nikomachischen Ethik* erläutert Aristoteles im so genannten Ergon-Argument (NE 1097b29[^]-1098a3), was gut für einen Menschen (aber auch für andere Lebewesen) ist. Diese Überlegungen haben die Kommentatoren sehr beschäftigt, denn sie scheinen ein Musterbeispiel vom pars-pro-toto-Fehlschluss darzustellen. Es wäre leicht zu bestimmen, was ein guter Mensch ist, sagt Aristoteles, wenn man nur wüsste, was seine Aufgabe oder spezifische Leistung sei. Er bemerkt, dass ein

¹ Von Wright fügt eine dritte Behauptung dazu, weil er auch die Frage diskutiert, ob Gesellschaften ein Wohl haben. Dies muss uns in diesem Kontext nicht interessieren.

guter Tischler seine Aufgabe gut erfüllt und ein gutes Auge seine Aufgabe gut erfüllt; dann fragt er, ob nicht der Mensch als solcher eine Aufgabe habe. Müssten wir angesichts der Tatsache, dass bestimmte Arten von Menschen (Tischler, Schuhmacher) eine Aufgabe haben und bestimmte Teile vom Menschen (Augen und Hände) eine Aufgabe haben, nicht annehmen, dass der Mensch selbst eine Aufgabe hat? Dann bestimmt er die Aufgabe bzw. die Funktion des Menschen:

Kann es sein, dass Tischler und Schuster Funktionen und Tätigkeiten haben und der Mensch keine hat, sondern faul geboren ist? So wie das Auge, die Hand, der Fuß und im Allgemeinen jedes Teil eine Funktion (*ergon*) zu haben scheint, könnte man annehmen, dass der Mensch auch eine Funktion neben allen diesen hat. Aber was wäre das? Denn das Leben scheint auch den Pflanzen gemeinsam zu sein, aber das Eigentümliche wird gesucht. Man muss also das ernährende und wachsende Leben beiseite tun. Es könnte ein wahrnehmendes Leben sein, aber dies scheint auch dem Pferd, dem Rind und jedem Tier gemeinsam. Es bleibt das tätige Leben dessen, der Vernunft hat. (NE 1097b29^{ff})

Zu seiner Entlastung soll angemerkt werden, dass Aristoteles seine eigene eher rhetorische Frage gar nicht beantwortet, sondern eine ganz andere Frage. (Genau genommen rettet er sich vor dem pars-pro-toto-Fehlschluß durch eine Äquivokation.) Er hatte gefragt, ob nicht der Mensch eine instrumentelle Funktion hat wie ein Handwerker oder ein Organ. Die Frage galt also, wozu oder wofür der Mensch gut sei. Die Antwort aber betrifft die Frage, was das Wohl des Menschen von dem Wohl der anderen Lebewesen unterscheidet. Die Ausführung der Funktion des Tischlers ist gut für andere – die Gemeinschaft; und die Ausführung der Funktion des Auges ist gut für ein anderes – den Menschen. Aber die Ausführung des *ergon* des Organismus (ob Mensch, Tier oder Pflanze) ist gut für den Organismus selbst. Aristoteles rekurriert wie selbstverständlich auf das Leben als das Bestimmende bei dieser Frage. Insofern stimmt er völlig mit von Wright überein; aber er weist auch auf den Grund hin, warum alles Lebendige ein Wohl hat: das *ergon*, das alles Lebendige mit den Pflanzen gemein hat ist Ernährung – oder anachronistisch: Stoffwechsel.

Der aristotelische Begriff des *ergon* zielt auf das, was ein Ding tut, das es zu dem macht, was es ist. Platons Rebschere und Aristoteles' Flötenspieler tun etwas, was sie zu dem macht, was sie sind. Durch Wachstum und Ernährung machen Pflanzen (aber alle anderen Lebewesen auch) sich zu dem, was sie sind – aber in einem ganz anderen Sinne. Das, was sie tun, gibt ihnen nicht nur einen Namen, wie Rebschere oder Flötenspieler, sondern macht sie buchstäblich und materiell zu dem was sie sind. Sie produzieren sich; und diese grundlegende organische Tätigkeit wird auch bei den höheren Formen Tier und Mensch vorausgesetzt.

Aristoteles selber realisierte keine dieser für uns sichtbaren Möglichkeiten, die in seinem Begriff des *ergon* steckten, da er den Begriff fest mit dem der Seele verbunden hatte. Aber in seiner Vorstellung von Ernährung und Wachstum ist eine genuin materielle Tätigkeit, die etwas zu dem machen kann, was es ist. Gerade diese Vorstellung einer selbst-produzierenden Tätigkeit wurde mit John Locke zur Grundlage moderner Organismustheorien in der Biologie (vgl. McLaughlin 2001).

Die charakteristische Tätigkeit eines Lebewesens ist gut für es selbst; die Ausführung der Funktion führt zur Erhaltung und Wiederherstellung des Funktionsträgers in seinem System; die Selbsterhaltung oder Selbstproduktion des Organismus ist seine charakteristische Tätigkeit. Gut für den Organismus und gut für die Tätigkeit der Erhaltung des Organismus sind fast dasselbe. Durch Stoffwechsel und Regeneration produziert ein Organismus seine eigenen Teile wieder. Der Beitrag eines Teils (seine Funktion) zu dieser Leistung des Systems führt zur Wiederherstellung aller Zellen, aus denen der Teil besteht – und kann so erklären, warum der Teil da ist und warum das System ein Wohl hat.

Dies ist natürlich noch lange keine ausgearbeitete Theorie der Wohlfahrt, keine erfolgreiche Naturalisierung des Wohls. Es bringt uns aber einen Schritt weiter als die Position von Wrights. Wir können jetzt sagen, unter welchen Bedingungen er Recht hatte, allen Organismen ein Wohl zuzuschreiben und allen Artefakten (auch dem Terminator) ein Wohl abzusprechen. Wer ein Wohl hat, dessen charakteristische Tätigkeit ist die Selbsterhaltung bzw. Selbstproduktion. Dies weist auch die Richtung auf, in die wir gehen müssen, wenn wir die beiden grundlegenden Probleme des ätiologischen Ansatzes lösen wollen und erklären wollen, ob Bewusstsein Funktion voraussetzt.

- Aristoteles (NE): *Die nikomachische Ethik*. Hrsg. von Rainer Nickel. Düsseldorf/Zürich 2001: Artemis und Winkler.
- Bigelow, John/Parfit, Robert (1987): Functions. In: *Journal of Philosophy* (84), 181–196.
- Cameron, James (1983): *Terminator Two. Judgment Day*. Screenplay unter <http://www.scifiscripts.com/scripts/t2.txt>.
- Cummins, Robert (1975): Functional Analysis. In: *Journal of Philosophy* (72), 741–764.
- Cummins, Robert (1983): *The Nature of Psychological Explanation*. Cambridge, Massachusetts: MIT-Press.
- Detel, Wolfgang (2001): Teleosemantik. Ein neuer Blick auf den Geist? In: *Deutsche Zeitschrift für Philosophie* (49), 465–491.
- Hempel, Carl Gustav (1959): The Logic of Functional Analysis. In: Gross, Llewellyn (Hg.): *Symposium on Sociological Theory*. New York: Harper and Row, 271–307. (Reprint in: Hempel, Carl Gustav (1965): *Aspects of Scientific Explanation*. New York: Free Press, 297–330.)
- Kant, Immanuel (1790): *Kritik der Urteilskraft*. Akademie-Ausgabe. Band V. Berlin/New York 1902ff.: de Gruyter.
- Kitcher, Philip (1993): Function and Design. In: *Midwest Studies in Philosophy* (18), 379–397.
- Macdonald, Graham/Papineau, David (Hg.) (2006): *Teleosemantics: New Philosophical Essays*. Oxford: Oxford University Press.
- McLaughlin, Peter (2001): *What Functions Explain*. Cambridge: Cambridge University Press.
- McLaughlin, Peter (2002): On Having a Function and Having a Good. In: *Analyse und Kritik* (24), 130–143.
- Millikan, Ruth (1984): *Language, Thought, and Other Biological Categories. New Foundations for Realism*. Cambridge, Massachusetts: MIT-Press.
- Millikan, Ruth (1989a): Biosemantics. In: *Journal of Philosophy* (86), 281–297.
- Nagel, Ernest (1961): *The Structure of Science*. New York: Harcourt Brace.
- Neander, Karen (1991): Functions as Selected Effects: The Conceptual Analyst's Defense. In: *Philosophy of Science* (58), 168–184.
- Neander, Karen (1995): Misrepresenting and Malfunctioning. In: *Philosophical Studies* (79), 109–141.
- Searle, John (1995): *The Construction of Social Reality*. New York: Simon and Schuster.
- Thomson, Judith J. (1997): The Right and the Good. In: *Journal of Philosophy* (94), 373–398.
- Wright, Georg Henrik von (1963): *The Varieties of Goodness*. London: Routledge.
- Wright, Larry (1973a): Functions. In: *Philosophical Review* (82), 139–168.
- Wright, Larry (1976): *Teleological Explanations. An Etiological Analysis of Goals and Functions*. Berkeley: University of California Press.