

## 5. GEIST

### 5.1 Leib und Seele

Platon läßt in seinem Dialog »Phaidon« den Sokrates beweisen, daß die Seele unsterblich sei – und Sokrates freut sich deswegen auf den bevorstehenden Tod durch den Giftbecher. Aristoteles dagegen identifiziert die Seele mit der Form des Leibes und betont, daß sie ohne einander nichts seien<sup>1</sup>: Die Diskussion über Leib und Seele ist so alt wie die Philosophie.

Eine besondere Schärfe bekommt diese Diskussion durch Descartes, der die ganze Wirklichkeit einteilt in zwei Substanzen, nämlich einerseits die denkende (res cogitans), die er in sich selbst als Subjekt seines Zweifels entdeckt, und die ausgedehnte (res extensa), die mit der gesamten materiellen Wirklichkeit auch den menschlichen Leib umfaßt; sie ist geometrisch, d.h. mathematisch zu behandeln.

Die moderne Naturwissenschaft und Technik hat dem „Leib-Seele-Problem“ im eben vergangenen Jahrhundert noch ganz neue Aspekte hinzugefügt und ihm eine außerordentliche öffentliche Aufmerksamkeit verschafft. Schon Konrad Lorenz hatte in seinem Institut in Seewiesen jahrzehntelang ein „Leib-Seele-Kolloquium“, das sich mit Problemen befaßte, auf die Lorenz und seine Mitarbeiter in der Diskussion ihrer verhaltensphysiologischen Ergebnisse stießen.

In neuerer Zeit ist die Diskussion durch zwei Entwicklungen wesentlich gefördert worden: Einerseits die atemberaubend rasche Entwicklung der Computertechnik in der zweiten Hälfte des 20. Jahrhunderts, angefangen mit dem gigantischen ENIAC (1946) mit 18.000 Röhren, der schon den Beinamen „Elektronengehirn“ bekam, bis zu heutigen PCs, deren jeder eine millionenfach größere Rechenleistung hat als der ENIAC, und von denen Millionen im Internet zusammengeschaltet werden können. – Trotzdem reicht die faktische Komplexität dieser gigantischen Ausrüstung noch immer nicht an die Komplexität eines einzigen menschlichen Gehirns heran<sup>2</sup>. Angesichts dieser Technik liegt die Frage nahe, die ich mit meinem Vergleich schon vorweggenommen habe, ob nicht ein genügend großer Computer genau das gleiche könnte wie ein menschliches Nervensystem, einschließlich aller Empfindungen, einschließlich Selbstbewußtsein – kurz, mit allem, was die menschliche Seele ausmacht.

Das zweite Gebiet, von dem die Leib-Seele-Diskussion ganz neu angeregt wird, ist die Neurophysiologie mit ihren faszinierenden Ergebnissen, die uns ganz neue Einblicke in die Arbeitsweise des Gehirns geben. Mikrotechnologie und bildgebende Verfahren ermöglichen uns, Vorgänge, die beim Denken ab-

<sup>1</sup> Aristoteles, de anima G5, 430a22-25.

<sup>2</sup> Auch wenn die Rechenleistung und Speicherkapazität aller Computer der Welt zusammen die eines menschlichen Gehirns übertrifft – ob das wahr ist, hängt bei der Verschiedenheit des Aufbaus sehr von der Vergleichsmethode ab –, sind die einzelnen Teile doch viel schwächer gekoppelt als es nötig wäre, um die Leistung eines Gehirns zu erreichen.

laufen, sehr genau räumlich und zeitlich im Gehirn zu verfolgen, und die Funktionsweise einzelner Nervenzellen lernen wir zunehmend besser zu verstehen. Auch hier muß man allerdings zugeben, daß diese sehr schönen Ergebnisse allenfalls ein Anfang der Erkenntnis davon, wie das Gehirn arbeitet, sein können.<sup>3</sup>

Der gesamte Wissenschaftsbereich, der sich mit diesen Fragen beschäftigt – von Computertheorie mit neuronalen Netzen, über Neurophysiologie, Neurologie und Psychiatrie, Verhaltens- und Evolutions-Biologie, Kognitionspsychologie und Linguistik bis zur Philosophy of Mind – wird unter dem neuen Namen *Kognitionswissenschaft* (“cognitive science”) zusammengefaßt.

## 5.2 ‚Mental‘ und ‚physisch‘

... in der Kognitions-  
wissenschaft

Die Frage, ob es gelingen kann, geistige Vorgänge vollständig auf materielle Vorgänge im Nervensystem zurückzuführen oder sogar im Computer zu simulieren – ob vielleicht das Nervensystem insgesamt „nichts als“ ein großer Computer ist –, hat philosophische Fragen angeregt, die sich vor allem damit befassen, was nun das besondere am Seelischen bzw. am *Mentalen* ist oder, was es überhaupt heißen kann, Seelisches auf Leibliches bzw. Mentales auf Physisches zurückzuführen<sup>4</sup>. Wir können hier aus diesem riesigen und schnell wachsenden Gebiet wiederum nur einige besonders markante Punkte herausgreifen und kommentieren, und verweisen im übrigen auf Übersichtsdarstellungen wie z.B. Beckermann und Metzinger oder die Textsammlung von Bieri<sup>5</sup>.

Eine große, vor allem  
angelsächsische Tradition Funktionalismus

In der Diskussion werden die verschiedensten Ismen unterschieden, wie etwa Materialismus, Physikalismus, Naturalismus etc., mit akribischen Fallunterscheidungen; dazu kann ich wiederum nur auf die Literatur verweisen. Ich möchte hier nur einen Argumentationsstrang näher beschreiben, der mir für die Entwicklung der Diskussion charakteristisch scheint, nämlich den *Funktionalismus*.

Identitätstheorie

In der Tradition des Wiener Kreises bzw. der analytischen Philosophie war zunächst die *Identitätstheorie* dominierend: Das Mentale ist mit dem entsprechenden Physischen identisch<sup>6</sup>. Eine solche Auffassung liegt wohl für die naturwissenschaftliche Denkweise nahe, so verstanden, daß mentale und physische Erscheinungen zu *demselben* Vorgang gehören können, zwei Aspekte derselben Sache sind<sup>7</sup>.

<sup>3</sup> Vgl. dazu z.B. Roth (1994).

<sup>4</sup> Da sich diese Diskussion in letzter Zeit vor allem im angelsächsisch dominierten Bereich der analytischen Philosophie abspielt, ist aus dem Leib-Seele-Problem das *mind-body-problem* geworden, rückübersetzt das Geist-Körper-Problem, das die Beziehung zwischen „mental“ und „physical“ betrifft, wobei letzteres auf deutsch, wie üblich, bald als *physikalisch*, bald als *physisch* zu übersetzen ist.

<sup>5</sup> Beckermann (1999), Metzinger (21996), Bieri (21993).

<sup>6</sup> Feigl (1958).

<sup>7</sup> Vgl. Weizsäcker (1977) S. 187-205.

Es zeigte sich bei näherer Analyse sehr schnell, daß die Rede von der *Identität* der beiden Erscheinungsformen einer schärferen logischen Analyse nicht standhielt. Jedenfalls wenn man, wie in dieser Tradition üblich, als Muster für Identität das Beispiel nimmt, daß *Scott* und der *Autor von Waverley* als Personen identisch seien, dann ist in diesem Sinn natürlich ein Gedanke, ich könnte jetzt vom Stuhl aufstehen, durchaus nicht identisch mit den neuronalen Entladungen im Gehirn, die ablaufen, wenn ich diesen Gedanken denke. Insbesondere könnte man dagegen einwenden, daß derselbe Gedanke, wenn Sie, der Leser, ihn denken, mit definitiv anderen neuronalen Vorgängen zusammenhängt, nämlich solchen in Ihrem Gehirn – obwohl es *derselbe* Gedanke ist.

Was heißt  
»Identität«?

Hilary Putnam<sup>8</sup> schlug als Lösung dieses Problems vor, daß für die physische Entsprechung des Gedankens ja nicht die konkrete Realisierung in den Nervenzellen des Denkers entscheidend sei, sondern die *Funktion* dieser Nervenzellen, die unabhängig davon ist, in welchem Gehirn sie realisiert („instantiiert“) ist. Putnam führte diesen Gedanken aus, indem er als Modell die Funktion eines Computers wählte, speziell des einfachsten möglichen Computers, der Turing-Maschine.

Die *Funktion* des  
Gehirns oder  
des Computers

### 5.3 Turingmaschine und Turingtest

Alan M. Turing hat schon sehr früh abstrakte Überlegungen zur Computertechnik angestellt, die heute noch häufig angeführt werden, unter den Namen *Turingmaschine* und *Turingtest*.

Die *Turingmaschine*<sup>9</sup> ist eine einfache, übersichtliche Einrichtung, mit der im Prinzip alle Funktionen ausgeführt werden können, die ein Computer, und sei er noch so kompliziert, ausführen kann. Sie besteht aus einem *Schreib-Lese-Kopf*, unter dem ein *Band* hindurchlaufen kann, und einer einfachen *Steuerung*. Man kann sich eine Art Morseapparat darunter vorstellen, der aber auf den Papierstreifen nicht nur schreibt, sondern auch von ihm lesen und Zeichen entfernen kann; vielleicht paßt die modernere Vorstellung von einem Kassettenrecorder besser, der an vordefinierten Stellen magnetische Marken auf das Band schreibt und von ihm liest. Die Vorrichtung arbeitet schrittweise: Bei jedem Schritt wird das Band um eine Stelle weitergeschoben, das dort befindliche Zeichen gelesen und ein neues Zeichen dorthin geschrieben. Die Stelle auf dem Band kann leer sein oder ein Zeichen enthalten. Man kann zeigen, daß ein Vorrat von zwei Zeichen genügt, z.B. 1 und 0. Außerdem ist die Steuerung des Apparates immer in einem Zustand; auch hier genügen zwei mögliche Zustände<sup>10</sup>, z.B. z1 und z2. Die Steuerung funktioniert nun nach einer „Maschinentafel“, aus der sie jeweils die nächste Aktion der Vorrichtung abliest. Und zwar besteht eine Aktion aus drei Teilen: a) ein Zeichen (0 oder 1) auf den Streifen zu schreiben, b) den Streifen entweder nach rechts oder nach links um einen Schritt weiterzuschieben, und c) die Steuerung in einen neuen Zustand (z1 oder

Die Turingmaschine –  
der einfachste Computer  
...

<sup>8</sup> Putnam (1960).

<sup>9</sup> Turing (1936).

<sup>10</sup> Vgl. Wiener (1998).

z2) zu bringen. Je nach dem augenblicklichen Zustand der Maschine gibt die Steuerung eine solche Aktion vor, abhängig davon, ob die Maschine vor der Aktion im Zustand z1 oder z2 ist, und welches Zeichen sie von dem Band gelesen hat. Diese Maschinentafel enthält also das *Programm* des Computers.<sup>11</sup>

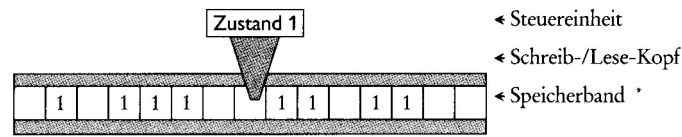


Abb 5.1: Schema einer Turingmaschine

... kann alles, was andere Computer können. Sind Computer intelligent?

Turing zeigt, daß es eine „universelle Turingmaschine“ gibt mit einer fest eingebauten Maschinentafel, der das Programm, soweit es über die Maschinentafel hinausgeht, und die Daten über das Band eingegeben werden. Er beweist mathematisch, daß eine solche einfache Maschine sämtliche Operationen ausführen kann, die ein Computer ausführt. Im Prinzip kann jedes normale Computerprogramm übersetzt werden in ein „Programm“ auf dem Streifen in der Turingmaschine; normalerweise wird allerdings jede einzelne Operation eines Computers in eine lange Reihe von Schritten der Turingmaschine übersetzt.

Der *Turing-Test* – der ebenfalls auf Alan Turing<sup>12</sup> zurückgeht – ist als Kriterium konstruiert, mit dem man entscheiden kann, ob ein Computer Intelligenz besitzt: Eine Testperson bekommt Ein- und Ausgabegeräte präsentiert (z.B. Tastatur und Bildschirm, es könnte aber auch Stimm-Ein- und Ausgabe sein), über die sie Fragen stellen und Antworten entgegennehmen kann. Diese Testperson soll nun mit geschickten Fragen herauskriegen, ob sich hinter der Ein-/Ausgabe ein Mensch oder ein Computer verbirgt. Wenn ein Computer auch von einem geschickten Tester nach intensiver Befragung nicht entlarvt wird, dann ist dieser Computer intelligent. – Turing hat 1950 vorausgesagt, daß innerhalb kürzester Zeit ein Computer seinen Test bestehen würde. Darin hat er sich geirrt. Inzwischen hat man vor der menschlichen Intelligenz viel mehr Hochachtung gelernt: Ein Computer wird schnell entlarvt, weil er sich „in der Welt nicht auskennt“, denn auch der schnellste Computer kann bisher nicht so viel „Kontextinformation“ ständig bereithalten, wie es ein Mensch ohne merkbare Anstrengung tut.

#### 5.4 Funktion

mathematischer Funktionsbegriff ...

Putnam dachte bei seiner Einführung der Funktion über das Modell der Turing-Maschine offenbar vor allem an das Programm des Computers bzw. eine Funktion im Sinne des mathematischen Funktionsbegriffs: Eine Vorschrift

<sup>11</sup> Eine schöne Einführung enthält man im Internet unter: „<http://www.matheprisma.uni-wuppertal.de/Module/Turing/>“.

<sup>12</sup> Turing (1950).

zum Abbilden bzw. Transformieren eines Maschinenzustands in einen anderen.

Eine der vielen Kritiken an Putnams Vorschlag lief darauf hinaus, daß dieser *mathematische* Funktionsbegriff gar nicht gemeint sein kann, wenn man von der Funktion des Nervensystems spricht, sondern daß dafür eher ein biologischer Funktionsbegriff in Frage kommt, wie etwa in dem beliebten Beispiel: Die Funktion des Herzens ist es, Blut zu pumpen. Dieser Funktionsbegriff hat, wie wir oben (4.16) gesehen haben, sehr eng zu tun mit *Zwecken* und *teleologischer* Betrachtung. Die Richtung in der Analytischen Philosophie, die diesen Funktionsbegriff bei der Analyse des Geist-Körper-Problems zugrunde legen will, nennt sich daher ausdrücklich „Teleo-Funktionalismus“. Dieser Ansatz gibt sicher das Problem bei der biologischen Beschreibung besser wieder als der mathematische Funktionsbegriff. Es zeigt sich aber, daß seine Behandlung im Rahmen der Analytischen Philosophie auf große Probleme stößt.

... oder teleologischer Funktionsbegriff?

Die Diskussion über diese neueren Ansätze wird z.Zt. noch sehr intensiv geführt<sup>13</sup>. Ich habe oben (4.16) schon erwähnt, daß ich bisher bei diesen Ansätzen dasselbe Problem sehe wie überhaupt bei der Behandlung der Teleologie im Rahmen der Analytischen Philosophie: Die Versuche, *Teleologie* – bzw. in diesem Fall *Funktion* – begrifflich auf Kausalität zurückzuführen, führen das ganze Unternehmen auf ein falsches Gleis.

Teleologie erklären durch Kausalität?

Die Debatte über das Leib-Seele-Problem ist seit Feigl und Putnam<sup>14</sup> so verzweigt und komplex geworden, daß ich nur paradigmatisch auf zwei vieldiskutierte Parabeln eingehen kann, die Chinesisch-Zelle und die Fledermaus:

### 5.5 Die Chinesisch-Zelle

John R. Searle ist u.a. ein Pionier der *Sprechakt-Theorie*. Diese Theorie knüpft an die Erkenntnis an, daß Sprache nur in besonderen Fällen zur Mitteilung objektiv vorliegender Tatsachen dient; in den meisten Fällen ist sie Bestandteil von anderen Sprach-Handlungen („speech acts“) wie etwa Warnen, Fragen, Wünschen, Trösten etc.

Sprachhandlungen

Im Rahmen dieser Überlegungen rückte eine Eigenschaft der Sprache ins Zentrum der Aufmerksamkeit, ihre *Intentionalität*. Damit ist gemeint, daß Sprache nicht einfach für sich besteht, sondern sich in bestimmter Weise auf die Wirklichkeit bezieht.<sup>15</sup> In diesem Sinne argumentiert Searle, daß ein Computer, der nur formale Operationen nach einem formal eingegebenen Programm vollführen kann, prinzipiell nicht in der Lage sei, menschliche Intelligenzleistungen zu vollbringen, z.B. Sprache zu verstehen. Searle schreibt dazu:

Intentionalität

“Suppose that I’m locked in a room and given a large batch of Chinese writing. Suppose furthermore (as is indeed the case) that I know no Chinese, either written or spoken, and that I’m not even confident that I could recognize

Die Chinesisch-Zelle

<sup>13</sup> Eine Übersicht gibt Stange (2001); für eine informationstheoretische Behandlung vgl. Lyre (2002).

<sup>14</sup> Feigl (1958), Putnam (1960).

<sup>15</sup> Searle (1983).

Chinese writing as Chinese writing distinct from, say, Japanese writing or meaningless squiggles.[...] Now suppose further that after this first batch of Chinese writing I am given a second batch of Chinese script together with a set of rules for correlating the second batch with the first batch. The rules are in English, and I understand these rules as well as any other native speaker of English. They enable me to correlate one set of formal symbols with another set of formal symbols, and all that “formal” means here is that I can identify the symbols entirely by their shapes. Now suppose also that I am given a third batch of Chinese symbols together with some instructions, again in English, that enable me to correlate elements of this third batch with the first two batches, and these rules instruct me to give back certain Chinese symbols with certain sorts of shapes in ‘response’ to certain sorts of shapes given me in the third batch.”<sup>16</sup>

Sprachverstehen  
im Kontext

Dies ist das berühmte “Chinese room”-Argument<sup>17</sup>. Searle bezieht sich bei seiner Fabel auf Ansätze aus der KI-Forschung zum Sprachverstehen (bzw. er karikiert solche Ansätze), bei denen der Kontext, in dem ein ausgesprochener Satz überhaupt nur einen Sinn hat, dadurch einbezogen wird, daß dem Computerprogramm eine bestimmte Szene vorgegeben wird. So erhält z.B. der Satz: „Bitte die Karte“, der zunächst alles mögliche heißen kann, sofort einen guten Sinn, wenn man weiß, daß ihn ein Mensch, der sich mit einer Begleiterin gerade in einem Restaurant hingesetzt hat, zum Kellner spricht. So sind die verschiedenen Papierstöße mit unverständlichen Zeichen und Handbücher mit verständlichen Verknüpfungsregeln zu verstehen, die Searle in seine Zelle bekommt, und nach denen er ebenso mechanisch verfahren kann wie ein Computer, der ein Programm der KI abarbeitet.

Verstehen im Gegen-  
satz zu formalem Ope-  
rieren

Searles Argument ist nun, daß ein Computer niemals in der Lage ist, chinesisch wirklich zu verstehen, genauso wenig wie Searle in der Zelle chinesisch versteht, wenn er nur ganz mechanisch („formal“) den Regeln aus seinem Handbuch folgt. Hier überspringt Searle freilich ein Problem, das in der späteren Diskussion eine große Rolle gespielt hat: Ich kann ja von mir nur sagen, *ich* verstehe (z.B.) englisch, nicht aber, mein *Gehirn* verstünde englisch. So muß man auch gar nicht erwarten, daß Searle, der in seiner Zelle sitzt und nach dem Handbuch unbekannte Zeichen verknüpft, die Sprache versteht, sondern allenfalls von dem gesamten System kann man das sagen. Was Searle vor allem hier vermißt, die Intentionalität, könnte nur durch entsprechende Eingabe- und Ausgabemöglichkeiten realisiert werden.

Searle erwidert auf diesen Einwand, ein Roboter mit Sensoren und Effektoren würde die Lage nicht verbessern, denn der müßte von einem Computer gesteuert sein, der ebenso formal operiere wie jeder Computer und daher keine Intentionalität haben könne<sup>18</sup>.

<sup>16</sup> Searle, (1980) S.417f.

<sup>17</sup> „Chinese room“ wird auf deutsch gewöhnlich als „Chinesisches Zimmer“ übersetzt, genauer müßte man vielleicht auf deutsch sagen „Chinesisch-Zelle“; im Englischen gibt es ja keinen Unterschied zwischen Adjektiv und Substantiv.

<sup>18</sup> Moor (1988), S. 41; Searle (1980).

Dieses Gedankenexperiment hat in den 20 Jahren seit seiner ersten Formulierung eine Flut von Literatur<sup>19</sup> ausgelöst, die sich vor allem auch mit der ungeheuer schwierigen Frage befaßt, was es eigentlich bedeutet, eine Sprache zu *verstehen*. Über die Frage von Verstehen und Bedeutung haben wir schon oben bei der biologischen Information gesprochen (4.8), wir werden unten bei der Frage nach der Reduktion darauf zurückkommen.

Verstehen und  
Bedeutung

## 5.6 Die Fledermaus

Einen weiteren Einwand gegen die Möglichkeit, menschlichen Geist mit Computern zu simulieren, möchte ich noch herausgreifen. Er ist mit einer ähnlich schönen und berühmten Fabel verbunden wie Searles Chinesisch-Zelle. In dieser Fabel wird gefragt: „Wie fühlt man sich als Fledermaus?“

Wie fühlt man sich als  
Fledermaus?

Thomas Nagel<sup>20</sup> diskutiert unter diesem Titel einen Aspekt des menschlichen Geistes, den man – so meint er – naturwissenschaftlich nicht erfassen und um so weniger im Computer simulieren könne, nämlich die *subjektive* Sicht. Die naturwissenschaftliche, objektive Beschreibung besteht aus *allgemeinen* Sätzen, ihre Objektivität besteht gerade darin, daß sie für jedermann zu jeder Zeit an jedem Ort gleichmäßig gilt. Das äußert sich z.B. darin, daß die Zeit in der Naturwissenschaft nur als ein Parameter *t* vorkommen kann, nicht aber in ihrer ganzen Struktur, in der es immer auch ein Jetzt gibt: Denn das Jetzt läßt sich nicht allgemein, im Sinn der Naturwissenschaft objektiv beschreiben.

die subjektive Sicht

So ähnlich, sagt Nagel, ist es mit *meinem* Blick auf die Welt, der gerade nicht allgemein und objektiv ist, sondern meiner und nur meiner. In diesem Sinn ist der Blick der Naturwissenschaft auf die Welt ein „Blick von nirgendwo“<sup>21</sup>. Aus diesem Grund können wir niemals dessen habhaft werden – so Nagels Fabel –, wie z.B. eine Fledermaus die Wirklichkeit sieht. Selbst wenn wir alles ausgemessen hätten, was in der Fledermaus vor sich geht, und uns „virtuell“ die Welt der Fledermaus genau simulieren könnten, dann könnte *ich* allenfalls die Welt sehen, wie *ich* sie sehe, wenn ich mich in die Umwelt einer Fledermaus hineinversetzte, aber eben nicht, wie eine Fledermaus *als Fledermaus* sie sieht. – Entsprechende Argumente gibt es auch in Bezug auf andere Menschen, z.B. die farbenblinde Forscherin Mary, die sich perfekt auskennt in der Physiologie und Neurophysiologie des Farbensehens und vollständig beschreiben kann, was im Menschen geschieht, wenn er rot sieht – trotzdem aber, da sie farbenblind ist, das *Erlebnis* des Rot-Sehens nicht haben kann. Das bedeutet – so der Einwand –, daß die perfekte naturwissenschaftliche Beschreibung bzw. die perfekte Computersimulation eben doch nicht dasselbe sein kann wie der menschliche Geist, da ihr dieses subjektive Moment fehlt<sup>22</sup>.

Das Problem  
der *Qualia*

<sup>19</sup> Vgl. z.B. – relativ willkürlich herausgegriffen – Moor (1988) und Boden (1983, 1989).

<sup>20</sup> Nagel (1974).

<sup>21</sup> Nagel (1986), IV.

<sup>22</sup> Dieses Argument der „Subjektiven Qualitäten“ erscheint in der Literatur unter dem Stichwort „Qualia“.

|   |  |
|---|--|
| „Ich“ und „Jetzt“<br>kommen als Inhalt der<br>Naturwissenschaft<br>nicht vor        | <p>Mir scheint, Nagels Beschreibung von Naturwissenschaft ist einfach richtig. Weder das Jetzt noch das Ich kommt in der Naturwissenschaft fundamental vor; sie ist so angelegt, daß es nicht vorkommen <i>darf</i>. Meines Erachtens ist das aber kein Einwand gegen die Möglichkeit einer naturwissenschaftlichen Beschreibung des menschlichen „Geistes“. Die naturwissenschaftliche Beschreibung soll ja gerade eine naturwissenschaftliche sein, also alle Bezüge auf Jetzt und Ich ausschließen. Das spricht aber nicht dagegen, daß sie trotzdem im Prinzip in der Lage sein könnte, das Arbeiten eines Nervensystems vollständig zu beschreiben und – wenn man einmal unbegrenzte technische Möglichkeiten annimmt – auch einen Menschen in einem Roboter zu simulieren. Ich würde vielmehr annehmen, daß ein solcher Roboter dann ebenfalls die Welt von seinem Standpunkt aus sieht, „Ich“ sagen und verstehen kann, und verstehen kann, was es heißt, wenn ich „Jetzt“ sage. Und wenn der Roboter Naturwissenschaft treibt, wird er auch einsehen können, daß diese beiden Begriffe in ihrem umfassenden Sinn in einer naturwissenschaftlichen Beschreibung nicht vorkommen können.</p> |
| Ein Computer ist ein<br>Werkzeug – mens-<br>chenähnlich braucht<br>er nicht zu sein | <p>Das bedeutet nicht, daß ich die Möglichkeit eines solchen Roboters behaupten will; allein die Komplexität, die er haben müßte, ist wahrscheinlich ein technisch unüberwindliches Hindernis. Vor allem aber ist nicht einzusehen, wozu es gut sein soll, eine so komplizierte Maschine zu bauen, während Millionen von wirklichen Menschen zu viel geringeren Kosten ohnehin herumlaufen. Da halte ich es lieber mit Winograd und Flores<sup>23</sup>, die in der Diskussion über künstliche Intelligenz den Aspekt betonen, daß ein Computer ein <i>Werkzeug</i> für die Tätigkeit von Menschen sein kann, und als solches für vieles sehr viel besser brauchbar als hilfsbereite Menschen sein könnte, daß er aber viele menschliche Fähigkeiten nicht hat und auch nicht braucht.</p>   |
| Ein großes Computer-<br>Programm ist wie eine<br>Bürokratie                         | <p>Ähnlich argumentiert Joseph Weizenbaum, selbst langjähriger prominenter KI-Forscher. Er stellt allerdings besonders heraus, wie sehr die verfügbaren Werkzeuge auch das bestimmen, was dann damit gemacht wird; und für die Computer insbesondere, daß wir die Programme, also unsere Werkzeuge, nicht mehr verstehen (schon 1976!): „Ein großes Programm ist ein kompliziert geknüpft Netz von Instanzen. [...] Die Formulierung eines Programms gleicht somit eher der Schöpfung einer Bürokratie als der Konstruktion einer [<i>mechanischen</i>] Maschine [...]“<sup>24</sup> – Auch damit nähert sich ein Computer menschlicher Intelligenz, jedenfalls in der Beziehung, daß er nicht – wie viele Geisteswissenschaftler argwöhnen – einem simplen Mechanismus gleicht.</p>   |
| Intelligente Computer<br>sind noch nicht am<br>Horizont                             | <p>Um es noch einmal zusammenzufassen: Faktisch können heutige Computer zwar sehr vieles viel besser als Menschen, aber bei vielen Fähigkeiten des menschlichen Geistes sind sie noch weit davon entfernt, auch nur Ansätze dazu zu liefern. Die Diskussion über die mögliche Simulierung des menschlichen Geistes durch Computer ist also eine Vorwegnahme von technischen</p>  |

---

<sup>23</sup> Winograd/Flores (1986).

<sup>24</sup> Weizenbaum (1977), S. 308.



Möglichkeiten, von denen niemand weiß, ob sie jemals realisiert werden. Andererseits sehe ich aber bisher keinen Grund, warum nicht prinzipiell, die technischen Möglichkeiten einmal unterstellt, alles am menschlichen Geist von Maschinen simulierbar sein soll. Die bisher vorgebrachten Argumente, das sei prinzipiell unmöglich, finde ich nicht überzeugend. Argumente, die sagen, es sei prinzipiell unsinnig, das zu versuchen, überzeugen mich dagegen viel mehr.

### 5.7 Freiheit

Die Frage nach der naturwissenschaftlichen Beschreibbarkeit spitzt sich besonders zu in dem Problem, wie ein Mensch *frei* entscheiden kann. Hier stoßen ganz verschiedene Aspekte zusammen: solche der Ethik, Moral, Rechtsphilosophie mit psychologischen, biologischen und neurophysiologischen Fragen bis hin zu unseren Problemen der Kognitionswissenschaft, der künstlichen Intelligenz und darüber schließlich der Mathematik und Physik.

Das Problem der „Willensfreiheit“ entsteht aus der folgenden Konstellation: Auf der einen Seite wissen wir, daß es freie und unfreie Entscheidungen gibt. Wenn ich einem Mitmenschen meine Brieftasche aushändige, weil er mich mit einer Pistole bedroht, fühle ich mich gezwungen, also unfrei; wenn ich dasselbe spontan tue, um diesem Mitmenschen eine Freude zu machen, fühle ich mich so frei wie nur möglich: ich hätte es auch lassen können, ohne einen Nachteil für mich befürchten zu müssen, und ich erweise mich sogar als frei von der Herrschaft des Mammon, indem ich nur zur Freude des anderen mein Geld verschenke. „Willensfreiheit“ ist der interessante Begriff deshalb, weil es natürlich immer Situationen geben kann, in denen ich nicht das tun kann, was ich will; aber das ist hier nicht die Frage. Die Grundlage jeder Frage nach Verantwortung ist vielmehr die Überzeugung, daß ich wenigstens dann, wenn ich etwas tue, weil *ich es will*, frei bin.

Auf der anderen Seite steht die Naturwissenschaft, die die Welt und in ihr die handelnden Menschen, objektiv nach Naturgesetzen beschreibt. Soweit die Naturgesetze auf den Menschen anwendbar sind, beschreiben sie sein Verhalten so objektiv wie den Lauf der Sterne; wenn also jemand die Bedingungen meines Handelns genau kennt und die naturgesetzlichen Zusammenhänge durchschaut, könnte er meine Handlungen voraussagen. Die klassische Formulierung dieser naturwissenschaftlichen Denkweise haben wir oben (1.4) bei Laplace gesehen. Danach wäre Freiheit nur Einbildung.

Das ist nicht nur ein Problem des klassischen mechanischen Determinismus: Die moderne Neurophysiologie hat begonnen, die Vorgänge im Gehirn beim Fassen eines Entschlusses zu beschreiben. Gerhard Roth<sup>25</sup> schreibt, daß bei einer Versuchsperson schon bevor sie sich bewußt ist, daß sie einen Entschluß faßt, im Gehirn Vorgänge ablaufen, welche die *Durchführung* des Entschlusses einleiten. Danach wäre also unser Bewußtsein, daß wir einen Ent-

Freiheit und  
Objektivität

Willensfreiheit

objektiv beschreibbar  
– also nicht frei?

Bewußte, freie Ent-  
scheidung: Nur eine  
Illusion?

<sup>25</sup> Roth (1994).

schluß fassen, erst die Folge davon, daß die Entscheidung irgendwo im Gehirn längst gefällt ist. Roth schließt daraus, daß Willensfreiheit nur Illusion sein kann. Hierzu möchte ich allerdings gleich anmerken, daß dieser Schluß nicht zwingend ist. Denn wir müssen ja nicht annehmen, daß das Ereignis im Gehirn, bei dem faktisch entschieden wird, welchen von zwei möglichen Wegen ich einschlagen werde, genau zu derselben Zeit stattfindet, in dem ich das Bewußtsein habe, mich zu entscheiden; in jedem Fall ist es *meine* Entscheidung. Es bleibt aber die Frage, wie und wann diese Entscheidung wirklich geschieht, wenn sie nicht unmittelbar mit meinem Bewußtsein davon zusammenfällt.

Gibt es einen Weg, diese beiden Seiten zusammen zu denken?

Kants Lösung

Kant hat eine Lösung im Rahmen seiner Transzendentalphilosophie vorgeschlagen, auf die ich kurz eingehen möchte, weil die kantische Sicht im Hinblick auf die Naturwissenschaft allgemein für uns höchst anregend ist.

Bedingungen der Möglichkeit von Erfahrung überhaupt

Nach Kant hat die Erfahrungswissenschaft ihre Strenge – ihre „Notwendigkeit und Allgemeinheit“, wie er sagt – daher, daß wir gar keine Erfahrungen machen könnten, wenn nicht gewisse Bedingungen erfüllt wären. *Daß* die »Bedingungen der Möglichkeit von Erfahrung überhaupt« erfüllt sind, wissen wir aber ohne spezielle Forschung, denn wir wissen ja, daß Erfahrung möglich ist, und das könnte nicht sein, wenn nicht die *Bedingungen* für ihre Möglichkeit erfüllt wären. Diese Bedingungen sind zugleich die Grundgesetze der Erfahrungswissenschaft. Das ist ein genialer Gedanke: Kant sagt, wir können die Dinge nur so beschreiben, wie sie uns erscheinen – eigentlich eine Trivialität. Wie sie uns erscheinen, hängt aber nicht nur von den Dingen an sich ab, sondern auch von der Art, wie wir Erfahrung machen. Darüber, wie wir Erfahrung machen, können wir aber etwas wissen, ohne daß wir uns dazu auf spezielle Erfahrungen berufen müssen, d.h. »a priori«. Wir brauchen dazu auch nicht psychologische oder ähnliche Kenntnisse, sondern schon die Analyse dessen, was wir überhaupt mit »Erfahrung« meinen, liefert uns die nötigen Einsichten.<sup>26</sup>

Kausalität (nur?) in der Erscheinung

Zu den Grundregeln der Erkenntnis, zu den »Bedingungen der Möglichkeit von Erfahrung überhaupt« gehört – für unser Beispiel maßgeblich – die kausale Verknüpfung: Die Welt kann uns nur *erscheinen*, d.h. Wirklichkeit für uns werden, wenn sie kausal verknüpft ist. Wir könnten Empfindungen nicht als Erfahrung einordnen, wenn sie nicht in einem Zusammenhang von Ursache und Wirkung stünden – so das Argument Kants. Diese Erkenntnis a priori erschließen wir, sagt Kant, aus der Art, wie *wir* Erfahrungen machen können, d.h., wie die Welt uns *erscheinen* kann; darüber, wie die Dinge an sich sein mögen, sagt uns das nichts.

Daß ich frei handle, kann nicht Erfahrung werden

In diesem Unterschied zwischen Erscheinung und »Ding an sich« sieht er nun auch eine Möglichkeit für die Freiheit<sup>27</sup>: Wer frei handelt, weiß das allenfalls von *sich selbst* als dem handelnden Subjekt, er ist in diesem Sinn ein »Ding an sich«. Er kann also, für sich

Daß ich frei handle, kann nicht Erscheinung werden

<sup>26</sup> Vgl. 3.14.

<sup>27</sup> Kant, KrV A444 / B472: Dritte Antinomie

selbst gesehen, frei handeln (d.h. eine neue Kausalkette beginnen), auch wenn sein Handeln sich in die Erfahrung *anderer* nur einordnen läßt als Teil einer Kausalkette; das ist Bedingung der Möglichkeit von Erfahrungen dieser anderen. Kant erinnert hier an die bekannte Unmöglichkeit, von einem Mitmenschen zu beurteilen, ob er frei gehandelt hat. Das kann man von außen nie sehen; man kann immer annehmen, daß es Ursachen gab, aus denen er nicht anders handeln konnte.

Der geniale kantische Gedanke der »Bedingungen der Möglichkeit von Erfahrung überhaupt« kann m.E. auch heute zum Verständnis der Naturwissenschaft beitragen, wie wir oben (3.14) gesehen haben. Ich bezweifle allerdings, daß er zugleich das Problem der Freiheit wirklich löst. Wir sind heute auch in einer ganz anderen Situation als Kant: Einerseits stellt sich uns das Problem viel schärfer, weil unsere Naturwissenschaft allumfassend geworden ist. Kant hätte sich wohl in seinen kühnsten Träumen nicht ausmalen können, wieviel man heute naturwissenschaftlich wirklich weiß – auch wenn das wiederum sehr wenig ist verglichen mit dem, was man noch wissen könnte. Die heutige Naturwissenschaft hat eine solche Einheit und Konsistenz erlangt, daß der Gedanke an ein Gesamtsystem von Naturgesetzen naheliegt, die alle aus wenigen fundamentalen Prinzipien folgen. – Auf der anderen Seite hat sich aber auch der Charakter der Wissenschaft grundlegend geändert: Kant hat noch als selbstverständlich den mechanischen Determinismus angenommen, wie Laplace ihn zehn Jahre nach Kants Tod klassisch formuliert hat. Heute lehrt uns die Quantenmechanik, daß man nicht annehmen kann, die Natur sei im Sinne der Laplace'schen Formulierungen determiniert. Eine solche Annahme führt zu Widersprüchen in der Naturwissenschaft selbst, die Quantenmechanik ist prinzipiell *indeterministisch*, der „Zufall“ ist in der Theorie unausweichlich enthalten.

Heute anders: Einheit der Physik und Indeterminismus

Pascual Jordan hat wohl als erster vorgeschlagen, nun in dieser Lücke die Freiheit anzusiedeln<sup>28</sup>: Was der Naturwissenschaft als Zufall erscheint, kann in Wirklichkeit das Ergebnis einer freien Entscheidung sein. Ich glaube allerdings, daß eine solche *Lücke* die Freiheit ebensowenig beschreiben kann wie das kantische »Ding an sich«.

Freiheit in der Lücke des Indeterminismus?

Dasselbe gilt für die Lücke ganz anderer Dimension, die Spaemann und Löw<sup>29</sup> für die Freiheit suchen: Das Subjekt, das autonom handelt, ist *Voraussetzung* dafür, daß es überhaupt Naturwissenschaften geben kann; dieses Subjekt, sagen Spaemann und Löw, könne doch nicht zugleich *Gegenstand* der Naturwissenschaft sein. Gerade von moralischer Entscheidung oder Verantwortung

Kann der Mensch, das Subjekt der Wissenschaft, zugleich ihr Objekt sein?

<sup>28</sup> Vgl. auch Jonas (1981).

<sup>29</sup> Spaemann/Löw (1985).

könne doch dort nicht die Rede sein, wo das Verhalten naturwissenschaftlich „erklärbar“ ist. Mit Vehemenz greifen daher diese Autoren vor allem die biologische Evolutionstheorie an und weisen den Formulierungen ihrer Vertreter ungerechtfertigte Schlüsse und Inkonsistenzen nach. Überwiegend haben sie Recht mit den Details ihrer Kritik, und ihr Anliegen teile ich ganz und gar. Trotzdem scheint mir, daß jeder auf verlorenem Posten kämpft, der innerhalb der Naturwissenschaft eine prinzipielle Lücke sucht als Refugium, sei es für Gott, Freiheit oder Unsterblichkeit. Die Naturwissenschaft hat regelmäßig Lücken solcher Art sehr bald gefüllt. Ihr prinzipieller Einwand läßt sich sogar relativ leicht entkräften: Der Mensch betreibt Wissenschaft, er ist Subjekt der Wissenschaft; aber *der Mensch* im allgemeinen muß nicht derselbe sein wie der, der gerade Wissenschaft betreibt. Die allgemeinen Gesetze, nach denen Menschen „funktionieren“, können auch für den Wissenschaftler gelten.

Universalität der Naturwissenschaft Denken Sie an die schöne Anekdote über Laplace, der zu Napoleon über den Schöpfer sagt: „Je n’avais pas besoin de cette hypothèse-là.“, (*Diese Hypothese habe ich nicht gebraucht*, vgl. 4.10). Der Laplaceschen Auffassung stimme ich zu. Die Naturwissenschaft ist universell in ihrem Anspruch, und ich zweifle nicht, daß alle Fragen, die überhaupt naturwissenschaftlich gestellt werden können, sich auch prinzipiell naturwissenschaftlich beantworten lassen.

Was ist Freiheit? Bleibt dann überhaupt noch Raum für Freiheit? Ich glaube ja, und um das zu sehen, müssen wir zuerst noch einmal betrachten, was Freiheit ist; d.h. wir müssen uns an die eigentlich philosophische Aufgabe wagen.

Freie Entscheidung: Nach Werten Eine freie Entscheidung bedeutet ja nicht eine zufällige oder willkürliche Entscheidung, sondern eine Entscheidung mit Gründen. Wenn mir zwei Möglichkeiten völlig gleichwertig erscheinen – sagen wir, ob ich auf mein Frühstücksbrot Honig oder Marmelade streichen will –, dann kann ich einem „zufälligen“ Augenblicksimpuls folgen oder auch eine Münze darüber werfen; aber das werde ich gerade nicht eine freie Entscheidung nennen, vielleicht überhaupt nicht eine *Entscheidung*. Frei vielmehr werde ich mich finden, wenn ich nach *Werten* entscheide, wenn ich das tue, was mir am meisten wert ist: Das kann die Freiheit sein, sich zu bereichern ohne staatliche Schranken, wie etwa im Manchester-Liberalismus oder unserer neuen globalen Ökonomie, oder die Freiheit der Selbstverwirklichung, bis hin zur „freien Fahrt für freie Bürger“ – oder es kann die Freiheit eines Janusz Korczak sein, sich umbringen zu lassen, obwohl er hätte davonkommen können, um des höheren Wertes der Solidarität mit „seinen“ Kindern willen.

Gerade freie Entscheidungen sind im Prinzip vorhersagbar ... Freilich weiß ich, daß man kaum je eine Entscheidung wird frei nennen können, daß ich oft genug etwas „will“, von dem ich weiß, daß ich es eigentlich nicht will. Aber in der Diskussion unseres abstrakten Beispiels der Willensfreiheit wird man relativ leicht zugeben, daß solche *natürlichen* Verhältnisse, welche die Freiheit behindern, auch naturwissenschaftlich beschreibbar sind. Das Problem entsteht beim *Ideal* einer absolut freien Entscheidung. Dafür gilt aber: Könnte ich die wirkliche Rangfolge der Werte eines Mitmenschen kennen,

dann könnte ich gerade bei seinen *freien* Entscheidungen das Ergebnis vorher-sagen. Und das bedeutet, daß auch freie Entscheidungen prinzipiell Gegen-stand von Naturwissenschaft – also von Gesetzen für Voraussagen – sein kön-nen.

Nach diesen Überlegungen besteht kein Widerspruch zwischen einer freien ... allerdings faktisch Entscheidung und der Möglichkeit, das Zustandekommen dieser Entscheidung zu komplex naturwissenschaftlich zu beschreiben – auch wenn faktisch eine solche Be-schreibung mit den Kenntnissen der heutigen Naturwissenschaft nicht vor-stellbar ist.

Heute ist allerdings der konsequente Gedanke einer naturwissenschaftlichen Beschreibung freier Entscheidungen eher denkbar als im Zeitalter des Mecha-nismus, der Gedanke nämlich, daß Vernunft und Verantwortungsbewußtsein, welche die freie Entscheidung leiten, ebenfalls zu den Möglichkeiten des Ge-hirns gehören, die naturwissenschaftlich beschreibbar sind. Wenn man die Po-sition nur konsequent genug zu Ende denkt, entfallen viele der traditionell dis-kutierten Probleme. Die heutige Wissen-schaft läßt die Willens-freiheit eher denken als die des 18. Jahrhunderts ...

Die Vorstellung einer naturwissenschaftlichen Beschreibung des Menschen war im 18. Jahrhundert vor allem deshalb so erschreckend, weil man als exakte Naturwissenschaft nur die Mechanik kannte und man damit den Menschen in Analogie zu den mechanischen Automaten denken mußte, die man im 18. Jahrhundert so kunstreich – wenn auch menschenunähnlich – verfertigte. Heu-te kennen wir eine breite Palette von Theorien, die prinzipiell die Einordnung auch von Vorgängen im Gehirn in die Naturwissenschaften ermöglichen. Wir erschrecken aber genauso wie unsere Vorfahren, wenn wir Leistungen des Verstandes vergleichen mit denen von Automaten, die wir heute als Analogie betrachten können, auch wenn diese Automaten nicht mehr mechanisch, son-dern elektronisch sind. Ein Automat kann fast nichts von dem, was ein Mensch kann. Wir wissen allerdings auch, daß ein Computer nach ganz anderen Prinzi-pien arbeitet als das Gehirn. Schon im Grad der Komplexität ist ein Computer mit einem Gehirn nicht vergleichbar. Man muß sich zudem vor Augen führen, daß das Leben eine Tradition von 3½ Milliarden Jahren hat, die Menschheit eine Geschichte von zigtausend Jahren: Daß wir diese Geschichte mit Kon-struktionen von heute im Labor ersetzen, ist kaum vorstellbar. ..., die vor allem aus Mechanik bestand

Das Erschrecken über solche falschen Parallelen kann aber in die Irre füh-ren, nämlich dann, wenn wir aus der Primitivität des naturwissenschaftlichen Modells auf die Primitivität des Erklärten schließen. Etwa wenn Spaemann und Löw die naturwissenschaftlich beschriebene Welt als „Welt der Faktizität“ be-schreiben – worüber man noch streiten könnte –, und dann unterstellen, „eine Welt der Faktizität kennt kein Sollen“<sup>30</sup>. Das wäre ja gerade eine *Frage* an die Naturwissenschaftler, wie man das, was hier „Sollen“ heißt, objektivierend be-schreibt. Von einer objektivierenden Beschreibung im Sinne der Naturwissen-schaft müßte man selbstverständlich verlangen, daß sie das wiedergibt, was wir Naturwissenschaft muß sich daran be-währen, daß sie die bekannten Phänome-ne rekonstruiert.

---

<sup>30</sup> Spaemann/Löw (1981), S. 259.

schon kennen, sonst ist sie falsch – eine naturwissenschaftliche Beschreibung müßte also das Sollen oder auch die Freiheit mitbeschreiben. Niemand weiß, wie weit man mit dem Versuch kommen würde, diese Begriffe objektivierend mit Inhalt zu füllen; aber niemand könnte auch heute behaupten zu wissen, daß das *nicht* geht.

Naturphilosophie muß  
die Tragweite der Wis-  
senschaft erörtern

Wir haben hier von der Naturwissenschaft das umfassendste, denkbar all-  
gemeinste Bild gemalt: das einer Theorie für alle Arten von objektiv prüf-  
baren Voraussagen. Mir scheint, daß man die Naturwissenschaft in ihrem umfassen-  
den Anspruch nur ernst nehmen kann, wenn man als möglichen Gegenstand  
auch das menschliche Denken und Handeln einbezieht, soweit es sich in objek-  
tiv prüf-  
baren Voraussagen fassen läßt. Zu den Aufgaben der Naturphilosophie  
gehört, die „Tragweite der Wissenschaft“<sup>31</sup> zu analysieren und sich über Be-  
hauptungen wie diese mit Gründen zu streiten.

Wenn wir den Bereich der objektivierenden Wissenschaft so weit ausdeh-  
nen, was bleibt dann überhaupt noch? Ist „in Wirklichkeit“ alles nur Physik  
und Chemie?

Auch die naturwissen-  
schaftlich beschriebene  
Freiheit ist Freiheit

Nein, denn es zeigt sich, daß der objektivierende Zugang der Naturwissen-  
schaft zur Wirklichkeit – zur *ganzen* Wirklichkeit – nur eine von vielen mögli-  
chen Zugangsweisen ist. Gerade bei unserem Beispiel, der Willensfreiheit, zeigt  
sich das sehr schön: Es ist durchaus vorstellbar, daß man die Struktur freier  
Entscheidungen objektivierend beschreiben kann. Das ändert aber nichts dar-  
an, daß *ich* mich für etwas *entscheide*, etwas will – und dieses Wollen ist auch ein  
Zugang zur Wirklichkeit, ein anderer als die objektive Beschreibung. Wir wer-  
den im nächsten Kapitel auf diese Fragen zurückkommen.

---

<sup>31</sup> Weizsäcker (1964/1990).

