

# Das Gehirn im Tank

## Vorwort

Im Wintersemester 2004/2005 besuchte ich (Charlie) im Rahmen meines Studiums die Vorlesung „Einführung in die Philosophie“ von Prof. Olaf Müller, wo er auf unterhaltsame Art und Weise bereits in der Einführungsveranstaltung das folgende Gedankenexperiment vorstellte. Mit dem interessanten philosophischen Beitrag „Wirklichkeit ohne Illusionen oder Der Abschied vom Skeptizismus“, der hier mit seiner ausdrücklichen Genehmigung veröffentlicht wird, gedenkt er zu beweisen, dass unsere Welt keine Illusion aus dem Simulationscomputer ist.

## Informationen zum Autor:

Prof. Dr. Olaf L. Müller, Jahrgang 1966, studierte Philosophie, Mathematik, Informatik und Volkswirtschaftslehre an der Universität Göttingen. Er wurde 1996 in Göttingen promoviert; die Habilitation folgte 2001. Forschungsaufenthalte an der University of California, Los Angeles und an der Harvard University. Olaf Müller unterrichtete Philosophie in Mannheim, Berlin (FU), Göttingen, Krakau und München (LMU). Seit Oktober 2003 ist er Professor für Philosophie mit dem Schwerpunkt Natur- und Wissenschaftsphilosophie an der Humboldt-Universität zu Berlin.

Webseite: <http://www.gehirnimtank.de>

## **Wirklichkeit ohne Illusionen**

**oder**

## **Der Abschied vom Skeptizismus**

**Seit den Matrix-Filmen streiten sich Informatiker und Laien, ob alle unsere Erlebnisse und Eindrücke der Welt vollständig auf Computersimulation beruhen könnten. Unser Autor behauptet, dass der Streit längst entschieden ist: durch einen völlig neuartigen Beweis, dessen geniale Grundidee der Harvard-Philosoph Hilary Putnam im Jahr 1981 publiziert hat. Wie sich zweifelsfrei demonstrieren lässt, ist die Welt garantiert keine Illusion aus dem Simulationscomputer.**

Klammheimlich hat gestern ein Bösewicht in Ihren Nachttrunk ein geschmacksneutrales Narkotikum gestreut, das direkt nach dem Einschlafen zu wirken begann. Sie waren allein zuhause, und so konnte niemand Alarm schlagen, als eine Bande von Chirurgen durch Ihr Schlafzimmerfenster einstieg, sich Ihres bewusstlosen Körpers bemächtigte und ihn in den Keller einer versteckten Klinik verschleppte. Die Chirurgen verloren keine Zeit. Sie sägten Ihren Schädel auf, um an Ihr Gehirn heranzukommen, das sie behutsam aus seiner Schale lösten und in eine Nährlösung gleiten ließen, damit es nicht abstirbt. Dann begann die Fummelarbeit.

Die Ärzte identifizierten jede einzelne Nervenbahn, durch die Ihr Gehirn bis gestern mit Ihrem Restkörper Informationen ausgetauscht hatte: Sehnerven, Nerven für akustische Reize aus dem Gehör, aber auch Nerven, durch die das Hirn Steuersignale zur Bewegung seines Ex-Körpers gesandt hatte. Alle diese (bei der Operation durchtrennten) Nervenstränge verbanden die Doktoren mit einem Computer, in den sie zuvor mit Akribie sämtliche Fakten über Ihr Haus, Ihre Familie, Ihren Job und Ihr Leben eingespeist hatten und in dem überdies ein geniales Programm zur Simulation von Nervenimpulsen geladen war. Als endlich die Wirkung des Narkotikums nachließ, starteten die Ärzte den Computer, und so meinten Sie, aus einem traumlosen Schlaf zu erwachen. Der Simulationscomputer sorgte zuverlässig für den Anschein von Normalität. Er simulierte das Strecken Ihrer Glieder, den Kälteschock unter der Dusche, den Geruch Ihres Morgenkaffees und die haptischen Qualitäten der Zeitschrift, die Sie jetzt in den Händen zu halten wähnen ...

Doch das simulierte Idyll trügt. In Wirklichkeit ist Ihnen von Ihrer gestrigen Existenz nur das Gehirn geblieben; es schwimmt in einem Tank mit Nährflüssigkeit herum. Und bei Ihnen zu Hause dampft nicht der Frühstückskaffee. Vielmehr durchstößt die Kripo Ihre Küche nach den Spuren der Entführer, und zwar genau jetzt!

Haben Sie irgendeine Chance herauszufinden, ob unsere kleine Geschichte erfunden ist? Können Sie wissen, dass Sie kein körperloses Gehirn im Tank sind, sondern dass Sie Hand und Fuß haben und soeben ein Druckerzeugnis aus echtem Papier lesen? Die Antwort auf diese Fragen ist offenbar negativ. Keine denkbare Beobachtung kann ausschließen, dass sie perfekt simuliert ist. Reden Sie sich nicht damit heraus, dass unsere Computer noch zu lahm wären, um die nötigen Simulationen in Echtzeit durchzurechnen. Denn wer garantiert Ihnen, ob

nicht irgendwelche Genies gerade gestern ihre geheime Arbeit an einem nie da gewesenen Super-Computer abgeschlossen haben? Unsere Geschichte ist theoretisch denkbar; das genügt.

Treiben wir die Sache auf die Spitze. Vielleicht ist Ihr Gehirn nicht erst gestern in den Tank geraten, sondern steckte von Anbeginn in dieser traurigen Lage? Dann hätten Sie niemals einen eigenen Körper gehabt. Und vielleicht befinden sich Tank, Hirn und Computer gar nicht im Keller einer irdischen Klinik. Sie könnten ja auch irgendwo im Andromedanebel herumschweben, weit, weit weg von unserem Sonnensystem. Wie, wenn es die Erde gar nicht gäbe? Dann wären Sie das einzige denkende Wesen überhaupt! - Wer denn in diesem Fall den Simulationscomputer programmiert haben soll? Nun, die ganze Konfiguration könnte durch einen gigantischen Zufall entstanden sein: unwahrscheinlich, zugegeben - aber theoretisch denkbar. Oder?

Nicht denkbar, behauptet der amerikanische Philosoph Hilary Putnam. Zur Begründung dieser Behauptung hat er einen raffinierten, verwirrenden und völlig neuartigen Beweis gefunden, um den sich die Fachwelt seit einem Vierteljahrhundert streitet. Der Streit ist leicht zu erklären. Hätte Putnam recht, so wäre dies eine philosophische Sensation historischen Ausmaßes. Seit Beginn der Neuzeit krankt die abendländische Philosophie daran, nicht verständlich machen zu können, dass wir irgend etwas über unsere äußere Umgebung wissen. Die Krankheit heißt cartesischer Skeptizismus, und sie verläuft so: Wir können nicht ausschließen, dass wir stets träumen oder andauernd von einem bösen Dämon getäuscht werden - oder dass wir seit jeher ein Gehirn im Tank sind. Also können wir auch nicht wissen, dass wir wirklich zwei Hände mit echten Fingern haben und genau jetzt Zeitungspapier zwischen diesen Fingern halten. Also können wir überhaupt nichts über die Welt um uns herum wissen. Diese skandalöse Konklusion wäre abgeschmettert, wenn Putnams Beweis gegen die Gehirne im Tank funktionierte (und wenn er sich auf die anderen beiden skeptischen Hypothesen, die von René Descartes alias Cartesius (1596-1650) stammen, übertragen ließe: auf ewige Träume und arglistige Dämonen).

Putnam braucht zwei sprachphilosophische Voraussetzungen, um seinen Beweis ins Rollen zu bringen. Die erste Voraussetzung ist eine Binsenweisheit:

(1) In meiner Sprache bezeichnet das Wort »Tiger« die Tiger.

Beachten Sie, dass dieser Auftakt des Beweises nicht viel Information über das Wort »Tiger« bietet. Aber das macht nichts. Je weniger ein Satz sagt, desto fester dürfen wir uns auf ihn verlassen:

Um zu wissen, dass Satz (1) zutreffen muss, braucht man keine Ahnung von Tigern zu haben und benötigt keinerlei empirisches Wissen über die Welt. Um der Wahrheit des Satzes sicher zu sein, braucht man lediglich eine besonders banale Information über die eigene Sprache. Man muss nur wissen, dass das Wort »Tiger« zur eigenen Sprache dazugehört (anders als beispielsweise das Wort »Lumpusch«). Dass das genügt, lässt sich anhand eines anderen Wortes demonstrieren:

Gesetzt, Sie wüssten, dass das Wort »Paläolepidopterologe« zu Ihrer Sprache gehört, seien sich aber nicht so sicher, ob das Wort die Erforscher versteinerner Schmetterlinge bezeichnet oder die versteinerten Schmetterlingsforscher (oder vielleicht noch etwas anderes?) Dann können Sie trotzdem richtig auf die folgende Frage antworten:

Was bezeichnet das Wort »Paläolepidopterologe«?

Sie streichen die Anführungszeichen am schwierigen Wort fort und sagen:

Das Wort bezeichnet die Paläolepidopterologen!

Damit haben Sie garantiert recht – obwohl Sie natürlich keine sonderlich informative Antwort riskiert haben. Und genauso steht es mit Putnams erster Voraussetzung:

(1) In meiner Sprache bezeichnet das Wort »Tiger« die Tiger.

Sie sagt wenig, gilt unter Garantie und kann daher nicht von den Skeptikern angegriffen werden. Gehen wir also weiter im Beweis. Putnams zweite Voraussetzung läuft so:

(2) In der Sprache eines ewigen Gehirns im Tank bezeichnet das Wort »Tiger« nicht die Tiger.

Warum nicht die Tiger? Laut Annahme hat ein ewiges Gehirn im Tank keinerlei kausalen Kontakt zu echten Tigern. Da ginge es nicht mit rechten Dingen zu, wenn das Gehirn die Tiger trotzdem bezeichnen könnte. Sein Gebrauch des Worts

»Tiger« müsste etwa durch irgendwelche geheimnisvollen nicht-kausalen Strahlen mit Tigern verbunden sein. Sprechen wäre eine übernatürliche Fähigkeit, eine Art Hexerei. Eine solche magische Auffassung von Sprache ist unhaltbar. Und wenn wir dieser magischen Auffassung widerstehen, dann akzeptieren wir eine - »externalistische« - Bedingung für Erfolg beim Bezeichnen: Ohne externe, kausale Verbindungen zu irgendwelchen Tigern kein Erfolg beim Bezeichnen mithilfe des Wortes »Tiger«.

Sie fragen, was denn dann Gehirne im Tank mit ihrem Wort »Tiger« bezeichnen, angesichts des mangelnden kausalen Kontakts zu echten Tigern? Vermutlich jene Bit-Muster im Simulationscomputer, die für die Tiger-Visionen verantwortlich sind. Wie auch immer diese Bits und Bytes aussehen, es sind keine Tiger aus Fleisch und Blut. Putnam mag sich in dieser Angelegenheit nicht genauer festlegen, und er muss es auch nicht. Meisterhaft beherrscht er die philosophische Kunst, gerade nur soviel zu sagen wie unbedingt nötig. Ihm genügt die negative Behauptung aus (2), der zufolge das Gehirn im Tank jedenfalls keine Tiger bezeichnen kann, weil das arme Gehirn, wie gesagt, von jedem Kausalkontakt zu Tigern abgeschnitten ist.

Nun beginnt die sensationell schnelle logische Arbeit. Aus den Voraussetzungen (1) und (2) ergibt sich zwingend:

(3) Meine Sprache ist von der Sprache eines ewigen Gehirns im Tank verschieden.

Das ist logisch; laut (1) und (2) haben die beiden Sprachen verschiedene Eigenschaften, also können sie nicht identisch sein. Wenn der Täter kleine Füße hat und der Gärtner große, dann war der Mörder nicht der Gärtner. Genauso bei Sprachen.

Und da sind wir schon am Ziel. Aus (3) folgt, und zwar abermals zwingend:

(4) Ich bin nicht seit jeher ein Gehirn im Tank.

Wäre ich nämlich ein ewiges Gehirn im Tank, so spräche ich auch seine Sprache, im Widerspruch zu Satz (3). Da (3) bereits bewiesen ist, kann ich nicht seit jeher ein Gehirn im Tank sein, Q.E.D.

Die Rekonstruktion des Beweises von Putnam, die ich Ihnen eben vorgeführt

habe, hat Crispin Wright aus St. Andrews (Schottland) gefunden. Sie ist einfacher zu überblicken als die ursprüngliche Fassung, die von Putnam stammt; zwar enthält sie immer noch einige Lücken, aber sie gibt Putnams Grundidee auf verblüffend einfache Weise wieder.

Wenige Zeitgenossen haben sich mit Putnams Beweis anfreunden können. Ich selber zum Beispiel hatte ganz zu Beginn den Verdacht, dass Putnams Beweis nichts ist als ein Taschenspielertrick. Andererseits war ich mir nicht sicher, ob ich den Beweis überhaupt verstehe. So habe ich mich entschlossen, der Sache auf den Grund zu gehen. Mithilfe der Deutschen Forschungsgemeinschaft (DFG) setzte ich im Herbst 1996 ein langjähriges Forschungsprojekt in Gang, in dessen Verlauf ich mit James Pryor und Hilary Putnam (Harvard), Holm Tetens (Freie Universität Berlin), Sven Rosenkranz (damals UNAM, Mexico City) sowie Wolfgang Carl, Felix Mühlhölzer und Thomas Schmidt (Göttingen) und Dominik Perler (Humboldt-Universität zu Berlin) kooperierte. Neben der Zusammenarbeit mit Fachkollegen habe ich in meinen Lehrveranstaltungen die unverbildeten Intuitionen anfangender und den Scharfsinn fortgeschrittener Studierender anzuzapfen versucht, zuletzt während einer erkenntnistheoretischen Vorlesung an der Humboldt-Universität im Sommersemester 2005.

Die Arbeit am Modell vom Gehirn im Tank und an Putnams Gegenbeweis brachte zwei Ergebnisse. Erstens war der Beweis in keiner der bislang bekannten Fassungen wasserdicht. Logische Lücken, versteckte Zirkularitäten und tausend Detail-Probleme verdarben den guten Eindruck, den der Beweis selbst bei Wohlwollenden hinterlassen hatte. In der Philosophie ist das keine Überraschung. Sogar den bedeutendsten Philosophen wird kaum je das Glück zuteil, endgültige Ergebnisse zu erreichen, gegen die sich nichts mehr sagen lässt. Fast immer sind die Kanonen der Kritiker stärker als die Panzer derer, die ein Problem lösen wollen.

Die Überraschung liegt im zweiten Ergebnis meiner Auseinandersetzung mit dem Beweis. Putnams Beweisidee ist völlig in Ordnung, ja sogar genial. Alle Makel des Beweises können ausgebügelt werden. So ist es mir gelungen, insgesamt drei wasserdichte Fassungen des Beweises auszuarbeiten, die sich gegen alle mir bekannten Einwände verteidigen lassen. Zwei dieser Einwände will ich jetzt vorführen und in aller Kürze entkräften. Laut erstem Einwand ist es misslich, den Beweis ausgerechnet mit dem Wort »Tiger« zu führen: Denn was wird aus dem schönen Gedankengang, wenn es zum Beispiel keine Tiger gibt? Dann könnte der

Beweis scheitern, weil sich meine Sprache dann doch nicht von der eines Gehirns im Tank zu unterscheiden braucht – wenn nämlich das Wort »Tiger« weder hier noch dort irgendetwas bezeichnen würde. (In der Diskussion mit den Skeptikern muss man stets das Äußerste befürchten. Immerhin haben sie uns ein Szenario vorgelegt, in dem es, abgesehen von der Tank-Konstellation samt Computer, überhaupt nichts gibt – insbesondere keinerlei Tiger).

Es ist sicherer, in dem Beweis anstelle von »Tiger« überall das Wort »Gehirn« zu verwenden. Was aber, wenn es keine Gehirne gibt? Einfach: Wenn es keine Gehirne gibt, sind wir ganz sicher kein Gehirn im Tank. Putnams Resultat (4) bleibt auch dann unangefochten.

Der zweite Einwand gegen Putnams Beweis besagt, dass der erhoffte Beweis in keiner denkbaren Fassung stimmen kann. Denn es liegt klar zutage, dass ein eingetanktes Gehirn den Beweis (in jeder beliebigen Fassung) wortwörtlich wiederholen könnte. Es könnte ganz genau dieselben Gedanken durchspielen, die wir uns zurechtgelegt haben; und diese Gedanken wären aus seiner Sicht genauso überzeugend, wie sie uns überzeugend vorgekommen sind. Aber das eingetankte Gehirn würde sich in einen fatalen Irrtum stürzen, wenn es unseren Beweis nacherzählen würde. Es käme, so wie wir, zu dem Schluss, nicht im Tank zu stecken – und das, obwohl es dem Tank sein Leben lang nicht entrinnen wird! Dürfen denn dann wir uns, endet der Einwand, auf ein Argument verlassen, das manche seiner Verwender (wie die Gehirne im Tank) so gefährlich auf Abwege lockt? Wäre das nicht leichtsinnig? Wer garantiert uns, dass wir nicht auch zu denen gehören, die von Putnams Beweis in die Irre geleitet worden sind und sich in falscher Sicherheit wiegen?

Die Antwort auf diese Frage beginnt mit dem Zugeständnis, dass die Gehirne im Tank den Beweis tatsächlich mit Recht wiederholen können. Denn ihr Ergebnis »Ich bin kein Gehirn im Tank« trifft vollkommen zu.

Wie das? Wie kann ein Gehirn im Tank die Wahrheit treffen, wenn es behauptet, gerade kein Gehirn im Tank zu sein? Einfach; wenn es sagt: »Ich bin kein Gehirn im Tank«, dann meint es mit diesen Worten überhaupt nicht, kein Gehirn im Tank zu sein. Denn seine Sprache funktioniert anders als unsere; nicht nur beim Wort »Tiger« (wie bereits dargelegt), sondern natürlich auch bei anderen Wörtern, etwa »Känguruh« oder »Eule«. Wenn das Gehirn im Tank »Tiger« beziehungsweise »Eule« sagt, dann meint es damit gewisse Konfigurationen aus

Bits und Bytes im Simulationscomputer - Bit-Tiger beziehungsweise Bit-Eulen. Genauso, wenn es »Gehirn« oder »Tank« sagt.

Nehmen wir also an, dass das eingetankte Gehirn mittels Putnams wunderbarem Beweis zu dem Ergebnis kommt: »Ich bin kein Gehirn im Tank«. Dann behauptet es (übersetzt in unsere Sprache), dass es kein Bit-Gehirn im Bit-Tank ist. Und damit hat es deshalb recht, weil es in der Tat kein Bit-Gehirn im Bit-Tank ist, sondern (gleichsam eine Ebene höher und weniger schlimm) bloß ein Gehirn im Tank (vgl. Abb. 2).

So gesehen, liefe Putnams Beweis auf diese Einsicht hinaus: Jedermann - Sie oder ich genauso wie das Gehirn im Tank - befindet sich exakt auf der Ebene, auf der er sich befindet; nicht etwa eine Ebene weiter unten.

Sie finden das trivial? Murren Sie nicht zu früh. So trivial Ihnen Putnams Beweis nach der neuesten Wendung unseres Gedankengangs auch vorkommen mag, der Beweis hat doch eine Konsequenz, die alles andere als trivial ist. Denn er führt uns vor Augen, dass die cartesischen Skeptiker einen grandiosen Fehler begehen, wenn sie unser Wissen über die äußere Welt leugnen, mithilfe von Szenarien wie dem vom Gehirn im Tank. Worin der grandiose Fehler besteht? Er besteht in dem verqueren Versuch, sich vorzustellen, dass wir in ein und demselben Moment woanders stehen als dort, wo wir gerade stehen.

Natürlich könnten wir woanders stehen als an unserem augenblicklichen Aufenthaltsort. (Wir könnten beispielsweise alle nach Feuerland durchgebrannt sein). Aber selbst dann bliebe der Satz »Ich bin, wo ich bin« wahr. Einerlei, ob ein Sprecher an die äußere Welt angeschlossen ist oder an einen Simulationscomputer oder an einen simulierten Simulationscomputer - genau dort, woher seine Sinneseindrücke stammen, liegen die Dinge, auf die sich sein Sprechen und Tun richten. Und wenn das Gehirn im Tank angesichts einer perfekten Hand-Simulation sagt: »Da ist eine Hand«, dann täuscht es sich damit nicht (anders, als es die Cartesier gern hätten). Denn dann redet es ja von einer Bit-Hand, und die ist im fraglichen Augenblick auch wirklich da. Schließlich verursacht sie den passenden handartigen Eindruck.

Es bedurfte eines philosophischen Genies vom Schlage Putnams, nach Jahrhunderten fruchtloser Debatten als erster zu sehen, dass dem cartesischen Skeptizismus mit solchen Überlegungen das Wasser abgegraben werden kann.



Hätten Sie irgendwelche Einwände? Dann senden Sie mir eine Flaschenpost aus Ihrem Tank.

Literatur:

Putnam, H.: Reason, truth and history, Kapitel 1 Cambridge: Cambridge University Press [Deutsch: Vernunft, Wahrheit und Geschichte (1982)] Frankfurt a. M. 1981.

Wright, C.: »On Putnam's proof that we are not brains in a vat«, In: Clark, Peter / Hale, Bob: Reading Putnam, 216-241, Cambridge 1994.

Müller, O. L.: Hilary Putnam und der Abschied vom Skeptizismus oder Warum die Welt keine Computersimulation ist. Wirklichkeit ohne Illusionen, Bd. 1, Paderborn 2003.

Müller, O. L.: Metaphysik und semantische Stabilität oder Was es heisst, nach höheren Wirklichkeiten zu fragen. Wirklichkeit ohne Illusionen, Bd. 2, Paderborn 2003.