

Das Leiden der Maschinen

Im Gespräch mit Thomas Metzinger

Herr Professor Metzinger, Sie sind einer der profiliertesten Philosophen und Bewusstseinsforscher weltweit. Was genau verstehen Sie unter Bewusstsein?

Die Frage nach dem Bewusstsein bezieht sich nicht auf ein einziges philosophisches Problem, sondern auf ein ganzes Bündel. Das Kernproblem beim Bewusstsein ist die Frage nach der Erste-Person-Perspektive, der Perspektive des sogenannten „phänomenalen Selbst“. In der Erfahrung des reinen Bewusstseins zum Beispiel existiert kein erkennendes Selbst. In dieser Form des unmittelbaren Erlebens handelt es sich eher um ein anonymes und unstrukturiertes Feld selbsterkennender Wachheit. Das normale phänomenale Bewusstsein, also das alltägliche Wachbewusstsein, das wir erleben, ist dagegen essenziell subjektiv, fast immer an eine personale Perspektive gebunden, und es galt deshalb in Wissenschaft und Philosophie lange als nicht verobjektivierbar. Auch deshalb habe ich bereits vor mehr als drei Jahrzehnten eine Theorie darüber entwickelt, wie in informationsverarbeitenden Systemen *ohne* ein Selbst eine robuste Form von Selbstbewusstsein und eine Erste-Person-Perspektive entstehen kann.

Viele Menschen haben die Intuition, dass das Bewusstsein ein essenziell biologisches Phänomen ist, das etwas mit der Lebendigkeit selbst – mit *Erleben* – zu tun hat. In der Philosophie spricht man diesbezüglich von Leiblichkeit oder *embodiment*, das heißt, dass Körper und Geist, also Denken und Fühlen, sich nicht nur wechselseitig bedingen, sondern letztlich zwei Aspekte eines einzigen, ganzheitlichen Phänomens sind.

Auf der anderen Seite ist es aber so, dass die weit überwiegende Zahl der Fachleute glaubt, dass nicht nur Intelligenz, sondern wahrscheinlich auch Bewusstsein „multirealisierbar“ ist, also unabhängig von der Hardware, von der rein physikalischen

Beschaffenheit des Systems. Die Unterscheidung zwischen natürlichen und künstlichen Systemen ist dabei – wenn man es philosophisch genau betrachtet – weder exklusiv noch erschöpfend. Wir werden es nämlich sehr bald mit Systemen zu tun haben, die weder künstlich noch natürlich sind. Aus Stammzellen gezüchtete, von Blutgefäßen durchzogene Zellklumpen zum Beispiel, sogenannte Organoide, die sich mithilfe von KI trainieren lassen. Intelligente Organoide sind keine

Maschinen, aber sie sind auch nicht natürlichen Ursprungs, weil sie nicht geboren wurden und nicht durch genetische Übertragung mit dem restlichen Strom des Lebens auf dem Planeten verbunden sind.

In Ihrer Theorie des Bewusstseins spielt das Selbstmodell eine wichtige Rolle. Was meinen Sie damit?

Es gibt kein Selbst als Substanz, das heißt keine vom Gehirn unabhängige ontologische Substanz, die sich eigenständig „im Sein halten“ könnte. Damit informationsverarbeitende Systeme wie wir ein robustes Ich-Gefühl bekommen, müssen sie ein Modell von sich selbst als einer Ganzheit entwickeln – in unserem Fall eine sogenannte transparente Repräsentation. Wenn wir zum Beispiel einen Apfel sehen, dann entsteht in unserem Gehirn die Repräsentation eines Apfels, die so gut und so schnell aktivierbar ist, dass wir sie gar nicht *als* Repräsentation erkennen. Philosophen nennen solche inneren Bilder „transparent“, weil wir sie nicht als Leistung unseres Gehirns bemerken, sondern sozusagen durch sie hindurchsehen, und dabei das Gefühl haben, wir wären direkt und unmittelbar mit der Welt in Kontakt. Wenn ein System ein solches transparentes Modell von sich selbst hat, also ein transparentes Selbstmodell, dann wird es das Erleben haben, sich *selbst* unendlich nah zu sein, sozusagen anhaftend, in direktem Kontakt mit dem Inhalt dieses Selbstmodells. Wenn Sie mich und Ihre Kollegen jetzt gerade auf dem Bildschirm sehen, sind Sie sich dessen bewusst, dass diese Bilder auf dem Bildschirm technisch erzeugt sind. Das nennt man „opake“ Repräsentationen. Das Paradebeispiel für solcherart erlebte Repräsentationen sind bewusst erlebte Gedanken. Von absichtlich initiierten Gedankenketten wissen wir ja, dass sie selbstkonstruiert sind. Bei Wahrnehmungen wissen wir das nicht. Die Einrichtung unseres Zimmers erscheint uns als direkt und unmittelbar gegeben. Ein auf diese Weise direkt und unmittelbar gegebenes Modell ist auch der eigene Körper, mit seinen Bauchgefühlen, Emotionen und unabsichtlichen Tagträumen. Diese Form der Repräsentation erzeugt die subjektive Identifikation in Form des „Ich-bin-das“-Gefühls. So entsteht auch das Gefühl der Meinigkeit: Das ist *mein* Arm.

Es ist deshalb nicht so, dass wir das Selbstmodell aus der subjektiven Innenperspektive als solches erleben würden. Der Begriff „Selbstmodell“ ist eine Beschreibung aus der Außenperspektive. Es ist sozusagen der Ursprung unserer subjektiven Perspektive, also das, was überhaupt erst eine Perspektive erzeugt.

der
blaue
reiter

*Es ist merkwürdig, wenn
Philosophen, die über
Bewusstsein schreiben,
keine eigenen
Erfahrungen mit
LSD machen wollen.*

Sie vertreten die These, dass es „neuronale Repräsentationen“ beziehungsweise „neuronale Korrelate des Bewusstseins“ gibt. Was meinen Sie damit?

Es gibt dazu ein englischsprachiges Buch aus dem Jahre 2000. Nehmen Sie zum Beispiel die Möglichkeit der Unterscheidung von feinsten Farbnuancen. Wir kön-



Prof. Dr. Thomas Metzinger,

Jahrgang 1958, studierte Philosophie, Religionswissenschaften und Ethnologie in Frankfurt am Main. 1999 erhielt er eine Professur für Philosophie der Kognition an der Universität Osnabrück. Von 1999 bis 2022 hatte er den Lehrstuhl für Theoretische Philosophie mit Schwerpunkten im 19. und/oder 20. Jahrhundert an der Johannes-Gutenberg-Universität Mainz inne. Als *Adjunct Fellow* war er Leiter der *MIND-Group* am *Frankfurt Institute for Advanced Study*, von 2005 bis 2007 Präsident der Gesellschaft für Kognitionswissenschaft und von 2009 bis 2011 Präsident der *Association for the Scientific Study of Consciousness*. 2022 wurde er in die Nationale Akademie der Wissenschaften (Leopoldina) gewählt.

Auszeichnungen (Auswahl):

- Leibniz-Vorlesungen an der Universität Hannover (2006)
- 2006–2007 Fellowship Embodied Communication Research Group
- 2008–2009 Fellowship Wissenschaftskolleg zu Berlin. Understanding the Brain
- 2014–2019 Fellowship Gutenberg Forschungskolleg
- 2021 Pufendorf-Medaille und The Pufendorf Lectures 2021 (Universität Lund, Schweden)

nen uns nur sehr schlecht an unterschiedliche Farben erinnern: Rot₃₁ und Rot₃₂ können wir diskriminieren, aber nicht zuverlässig re-identifizieren. Aber wenn wir aus einer großen Anzahl verschiedener Farbkarten das reinste Blau, das reinste Gelb, das reinste Rot und das reinste Grün auswählen sollen, dann suchen alle Menschen dieselben Karten aus. Weil das etwas mit der Funktionsweise unseres Gehirns zu tun hat, können wir es wissenschaftlich untersuchen. Die Frage ist deshalb, ob es auf der physikalischen Ebene ein minimal hinreichendes Korrelat gibt, das heißt einen messbaren Zustand unseres Gehirns, der mit der Wahrnehmung von zum Beispiel Rot₃₁ korrespondiert oder womöglich sogar damit identisch ist. Es geht dabei nicht um eine reduktionistische Ideologie. Aber es ist im letzten Jahrhundert sehr klar geworden, dass die Inhalte des bewussten Erlebens stark „von unten“, das heißt physikalisch determiniert sind. Man kann zum Beispiel mit Elektroden im Gehirn bestimmte Handlungen auslösen und bestimmte Gefühle erzeugen. Außerdem können wir sehr viele Dinge bewusst erleben, die wir sprachlich nicht beschreiben können. Man kann zum Beispiel einem Blinden nicht wirklich mitteilen, was Rot ist und es gibt auch Menschen, die ganz andere Farben sehen als der Rest der Menschheit. Es könnte auch sein, dass zum Beispiel die Unterscheidung von Farben gar keine physische Eigenschaft des Gehirns ist, dass es also kein rein neuronales Korrelat gibt, sondern eher eine Art, wie das Gehirn rechnet, eine *funktionale* Eigenschaft, eine bestimmte Art wie Information fließt.

Würden Sie die Willensfreiheit bestreiten?

Eine meiner größten Leistungen besteht darin, dass ich mich aus der deutschen Debatte um die Willensfreiheit erfolgreich ferngehalten habe. Viele Menschen können und wollen sich nicht vorstellen, dass es keine Willensfreiheit in einem starken Sinne gibt. Sie würden sich wie Schizophrene fühlen, die das Gefühl haben, dass ihre Gedanken und Körperbewegungen nicht die eigenen sind. Gerade die Altlinken aus meiner Generation und davor halten das für brandgefährlich, weil fehlende Willensfreiheit angeblich das revolutionäre Subjekt in Frage stellt. Es gibt also mögliche Erkenntnisse, die einfach nicht wahr sein dürfen, weil man sie für politisch gefährlich hält oder weil sie uns weh tun, sich intuitiv als „krank“ anfühlen.

Aber jede Handlungsentscheidung, die wir bewusst erleben, hat einen unbewussten Vorläufer in der Informationsverarbeitung im Gehirn. Das heißt, dass im Gehirn vorher etwas passiert, das wir nicht erleben können. Wenn wir Fahrrad fahren lernen, dann machen wir das bewusst. Aber das Zeichen dafür, dass wir es wirklich gut können, ist, dass wir unbewusst fahren und nicht mehr auf unser Gleichgewicht achten müssen. Wenn wir den überraschenden Impuls haben, nach etwas zu greifen, dann muss unser bewusstes Selbstmodell diese Überraschung „wegerklären“, denn die Vorläuferprozesse zu dieser Handlung können wir nicht bewusst erleben. Dieser Moment ist die

Abbildung:
Thomas Metzinger

Mit freundlicher
Genehmigung der
Onassis Foundation,
Athen.