

Julian Nida-Rümelin

Die Freiheit des Geistes

(Arbeitstitel)

Textbaustein für den geplanten gemeinsamen FAZ-Artikel mit David Linden

Der Disput zwischen Peter Janich und Wolf Singer endete schließlich unbefriedigend (FAZ vom ...). Singer forderte Janich auf, ein Experiment vorzuschlagen, das die Unabhängigkeit des Mentalen zeige, Janich antwortete darauf indem er aus wissenschaftstheoretischen Überlegungen heraus die Möglichkeit eines solchen *experimentum crucis* bestritt. Zweifellos ist es naiv anzunehmen, man könne diese grundlegende Frage des Verhältnisses von Geist und Materie mit einem Experiment klären, aber andererseits sollte die behauptete Nicht-Reduzierbarkeit des Geistigen, sprechen wir hier ruhig plakativ von der „Freiheit des Geistes“ sich durch empirische Befunde stützen lassen. Dazu nehmen wir im Folgenden Stellung: Lassen sich empirische Befunde denken oder gar experimentelle Anordnungen entwickeln, die für die Freiheit des Geistes sprechen? Der eine von uns ist von der Freiheit des Geistes überzeugt, er definiert menschliche Freiheit über das kausale Wirken von Gründen, die ihrerseits nicht naturalisierbar sind (vgl. Julian Nida-Rümelin: Über menschliche Freiheit, Stuttgart 2005). Der andere hat zu dieser philosophischen Frage keine abgeschlossene Meinung, er ist experimentell arbeitender Neurowissenschaftler, der aufgrund experimenteller Befunde die Annahme einer gewissen Unabhängigkeit des Mentalen für plausibel hält. Wir stellen vier Experimente vor und jedes hat den Charakter eines Gedankenexperimentes, von oben nach unten zunehmend sind jedoch konkrete experimentelle Anordnungen denkbar, die heute schon oder jedenfalls in der Zukunft eine empirische Überprüfung möglich machen.

1. Eine Person bildet sich eine Meinung über den Zustand ihres Gehirns. Diese Meinung realisiert sich – wie Singer und wir gemeinsam annehmen – in einem neurophysiologischen Prozess bzw. in einer Eigenschaft des neurophysiologischen Zustandes des Gehirns. Wenn der Zustand z des Gehirns die neurophysiologische Realisierung der Überzeugung, dass das Gehirn im Zustand z dies einschließt, dann ist diese Überzeugung wahr. Zugleich ist dann (zum gleichen Zeitpunkt, kontrafaktisch) die Überzeugung der Person wahr, dass sein Gehirn

nicht im Zustand z ist, weil diese Meinung sich in einem anderen Gesamtzustand seines Gehirns z' niederschlagen würde. In Singer'scher Terminologie haben wir es hier mit einem Paradoxon zu tun, das sich aus der Selbstbezüglichkeit des Gehirns ergibt. Die Identifikation von Überzeugungen mit neurophysiologischen Zuständen führt in einen Widerspruch.

2. Kausalrelationen haben, jedenfalls nach der klassischen Regularitätstheorie der Kausalität, einen algorithmischen Charakter, d.h. aus einem gegebenen Zustand ergibt sich auf Grund naturwissenschaftlicher Notwendigkeiten zwingend ein einziger Nachfolgezustand. Singer hat mehrfach deutlich gemacht, dass er dieses Verständnis von Kausalität auch für das neurophysiologische System zugrunde legt. Viele unserer Überzeugungen können jedoch nicht algorithmisch begründet werden. Schon die Theoreme der Prädikatenlogik erster Stufe lassen sich algorithmisch nicht beweisen. Wenn Begründungen nichts anderes wären, als neurophysiologische Prozesse und diese ihrerseits deterministischer Kausalität gehorchten, dann wäre die menschliche Fähigkeit auch komplexere Überzeugungen zu begründen, ein Rätsel.

3. Auch wenn man über Singer hinausgehend, ein probabilistisches Verständnis von Kausalität in die Betrachtung mit einbezieht, wird die Sache nicht besser. Eine rationale Person P hat zwei Behälter vor sich: der eine Behälter ist durchsichtig und enthält offensichtlich einen Tausendeuroschein, der andere Behälter ist nicht durchsichtig, aber P weiß, dass er entweder nichts oder eine Million Euro enthält. P steht vor der Alternative, entweder den Inhalt beider Behälter an sich zu nehmen oder sich auf den Inhalt des undurchsichtigen Behälters zu beschränken. Da P rational ist, wird sie den Inhalt beider Behälter an sich nehmen, denn unabhängig davon, ob der undurchsichtige Behälter eine Million Euro enthält oder nicht, sichert P sich dadurch zusätzlich tausend Euro.

Es scheint völlig unerheblich zu sein, auf welche Weise der Inhalt des undurchsichtigen Behälters bestimmt wird. Entscheidend ist nur, dass vor der Entscheidung von P und unabhängig von der Entscheidung von P feststeht, ob der Behälter eine Million Euro enthält oder nicht. Man stelle sich nun vor, dass eine Person (V), die mit der Persönlichkeit und den Handlungsweisen von P vertraut ist und ihr zukünftiges Verhalten in zahlreichen Fällen zutreffend vorausgesagt hat, über den Inhalt des undurchsichtigen Behälters bestimmt. P weiß, dass V die eine Million in den betreffenden Behälter genau dann legt, wenn V vorausgesagt hat, dass P nur den Inhalt des undurchsichtigen Behälters an sich nimmt. Wenn man nun zusätzlich annimmt, dass V 's Voraussage in Bezug auf P 's Verhalten so gut wie immer zutrifft, dann gibt es eine hohe Wahrscheinlichkeit dafür, dass V die eine Million

hinterlegt hat, wenn P sich dafür entscheidet, nur den Inhalt des undurchsichtigen Behälters an sich zu nehmen; denn die hohe Wahrscheinlichkeit einer richtigen Vorhersage äußert sich darin, dass die Entscheidung von P einerseits und die Entscheidung von V, eine Million zu deponieren andererseits nicht probabilistisch unabhängig sind, obwohl das Beispiel so gewählt ist, dass eine kausale Beeinflussung ausgeschlossen ist.

Eine Lösungsstrategie des Newcomb'schen Problems besteht darin, die im Erwartungswert-Maximierungskriterium zulässigen subjektiven Wahrscheinlichkeitsfunktionen einzuschränken. Diese Lösungsstrategie hat zur Entwicklung der sog. kausalen Entscheidungstheorie geführt. Eine Beschränkung dieser Art hat das Ziel, den Konflikt des Dominanzkriteriums mit dem Erwartungswert-Maximierungskriterium auszuschließen. Die Übereinstimmung von Dominanzkriterium und Erwartungswert-Maximierungskriterium führt zwar im Newcomb-Beispiel zu der intuitiv überzeugenden Empfehlung, den Inhalt beider Behälter an sich zu nehmen, hat aber in strukturell ähnlichen Interaktionssituationen wie z.B. dem Gefangenendilemma zur Folge, dass eine rationale Person nicht kooperativ entscheiden kann: beim Gefangenendilemma führen die dominanten individuellen Strategien zu einem Ergebnis, das nicht pareto-optimal ist. Eine genauere Betrachtung der Newcomb-Situation zeigt, dass implizit von einer kausalistischen Auffassung menschlichen Handelns Gebrauch gemacht wird. Denn die probabilistische Abhängigkeit der beiden Ereignisse (was wählt P und was tut V) setzt voraus, dass es so jemanden wie V geben könnte. Das Newcomb-Problem entsteht jedoch erst dann, wenn V's Prognosen über das Entscheidungsverhalten von P für beliebige Entscheidungssituationen zuverlässig sind. Das Newcomb-Problem kombiniert in genialer Weise zwei scheinbar sichere Intuitionen: die Intuition, dass menschliches Handeln kausalen Determinanten (und seien sie probabilistisch) unterliegt und daher, soweit erklärbar, grundsätzlich auch prognostizierbar ist und andererseits die Intuition, dass es Situationen gibt, in denen eine Person zwischen Alternativen frei wählen kann. Das Beispiel ist so konstruiert, dass diese beiden Intuitionen zugleich nicht aufrechterhalten werden können. Newcombs Problem ist ein Argument gegen die Vorhersehbarkeit der Handlungen einer rationalen Person: P sollte den Inhalt beider Behältnisse an sich nehmen¹.

4. Wenn es möglich wäre Experimente nachzuweisen, dass menschliche Intentionen eine kausale Rolle für neurophysiologische Prozesse im Gehirn spielen, dann wäre ganz

¹ Weitere Gesichtspunkte des Newcomb-Paradoxons behandeln: T.M. Benditt/D.J. Ross, „Newcomb's Paradoxon“, *Brit J. Phil. Sc.* 27 (1976) S. 161-164; W. Spohn, *Grundmodelle der Entscheidungstheorie*, Kronberg/Ts. 1978, Kap. 5.1 und 5.2; R.M. Sainsburg, *Paradoxes*, Cambridge 1988, Kap. 3.

unabhängig davon, ob die bisherigen Gedankenexperimente akzeptiert werden, der von Singer geforderte Nachweis erbracht. Eine solche experimentelle Anordnung müsste die genetische und epigenetische Ausstattung, sowie die Vorgeschichte sensorischer Stimuli konstant halten und in einer gegebenen Situation lediglich die Intentionen variieren. Wenn diese Variationen der Intention selbst (ohne eine Veränderung der sensorischen Stimuli) Veränderungen im neurophysiologischen System nach sich ziehen, dann wäre gezeigt, dass Intentionen kausal relevant sind. In der Tat ist dies die naheliegendste Interpretation der Fähigkeit von Menschen und Primaten über ein Neuro-Implantat vermittelt etwa den Cursor eines Computers zu steuern. Die experimentelle Anordnung müsste die Person etwa in einen völlig abgedunkelten Raum ohne Geräusche setzen, so dass sicher gestellt ist, dass die sensorischen Stimuli konstant bleiben (die Vorgeschichte sensorischer Stimuli ist ex definitione konstant, ebenso die genetische und epigenetische Ausstattung) und die Person diktiert lediglich ihre Absichten auf Band. Gleichzeitig wird untersucht, ob mit diesen Intentionen neurophysiologische Prozesse korrelieren.

5. Entkoppelung des Mentalen vom Neurophysiologischen (Linden)

Anmerkung: Dieser Artikel ist aufgrund eines Gespräches entstanden, das David Linden und Julian Nida-Rümelin im Rahmen einer Tagung in der Villa Vigoni zum Thema „Identität: Vergessen und Gedächtnis“ hatten, die neurowissenschaftliche und philosophische Perspektiven zusammenführte.