

Kapitel V

Der Nutzen von Befriedigung: Ein Zirkel in der Entscheidungstheorie¹

ANNA KUSSER UND WOLFGANG SPOHN

1. Eine kausale Voraussetzung der Entscheidungstheorie
2. Eine Verletzung dieser Voraussetzung
3. Wie man dem Zirkel nicht entgeht
4. Wie man dem Zirkel gerecht werden kann

Niemand wird bezweifeln, daß wir intrinsische Wünsche haben. Denn sicherlich haben wir Wünsche, und viele von ihnen sind abgeleitet. Demnach müssen – wie bei einer deduktiven Theorie, deren beweisbare Aussagen in Axiome und Theoreme zerfallen – einige unserer Wünsche grundlegend sein.

Weit weniger klar ist dagegen, worauf sich unsere intrinsischen Wünsche beziehen oder beziehen sollten. Unbestreitbar dürfte immerhin sein, daß neben anderem Glück, Freude, Schmerzfreiheit, gewisse Empfindungsqualitäten, Gefühle und Stimmungen, eben allgemein die eigenen psychischen Befindlichkeiten, Gegenstand intrinsischer Wünsche sind; die Wünsche danach scheinen einer weiteren Begründung weder fähig noch bedürftig.

Wünsche sind mehr oder weniger stark, und sie bestimmen, zusammen mit mehr oder weniger festen Überzeugungen, rationales Handeln. Den unterschiedlich starken Wünschen und Überzeugungen entsprechen im

¹ Diese Arbeit ist eine Übersetzung und leichte Überarbeitung unseres Aufsatzes 'The Utility of Pleasure is a Pain for Decision Theory', *Journal of Philosophy* 89 (1992), S. 10–29; sie erscheint hier mit freundlicher Genehmigung des *Journal of Philosophy*. Wir sind sehr dankbar für die Gelegenheiten zur Diskussion dieses Aufsatzes, die uns an der Freien Universität Berlin, an der Universität von Tilburg, der University of California at Irvine und der Universität des Saarlandes geboten wurden. Insbesondere danken wir Jack Birner, David Gauthier, Bert Hamminga, Dan Hausman, Isaac Levi, Gregory Kavka, Karel Lambert, Alan Nelson, Rainer Trapp und Ernst Tugendhat für hilfreiche Kommentare. Die Arbeit wurde teilweise durch die Deutsche Forschungsgemeinschaft unter dem Vorhaben Nr. Sp 279/2–1 unterstützt.

formalen Rahmen der Entscheidungstheorie die subjektiven Nutzenwerte und Wahrscheinlichkeiten. Auch die Unterscheidung zwischen intrinsischen und extrinsischen Wünschen findet in der Entscheidungstheorie ihren Niederschlag, da jedes Modell eine – jedenfalls relativ zum jeweiligen Modell – grundlegende Nutzenfunktion enthält sowie eine Funktion des erwarteten Nutzens, die sich aus der grundlegenden Nutzenfunktion und der subjektiven Wahrscheinlichkeitsfunktion herleitet.² In der Tat kann die Entscheidungstheorie für sich beanspruchen, die soweit insgesamt beste formale Explikation der intuitiven, von Wünschen und Überzeugungen ausgehenden praktischen Überlegung zu liefern. Mithin ist davon auszugehen, daß die intrinsisch gewünschten psychischen Zustände einer Person im Definitionsbereich ihrer Nutzenfunktion liegen. Dann aber, so behaupten wir, versagt die Entscheidungstheorie. Genauer gesagt, wollen wir zwei Thesen verteidigen:

- (A) Wenn Entscheidungssituationen eine gewisse kausale Struktur aufweisen, kommt die Entscheidungstheorie in Schwierigkeiten, da sich in ihnen kein erwarteter Nutzen bestimmen läßt.
- (B) Entscheidungssituationen dieser kausalen Struktur liegen tatsächlich häufig vor, nämlich in der Regel dann, wenn die Nutzenfunktion für intrinsisch evaluierte psychische Zustände definiert ist.

Der Abschnitt 1 erläutert die These (A); im Abschnitt 2 wird die These (B) verteidigt. Dem Verdacht, daß die aufgewiesene Unzulänglichkeit der Entscheidungstheorie auf irgendwelchen Mißverständnissen beruht, versuchen wir im Abschnitt 3 zu begegnen. Obwohl wir für unser Problem keine Lösung haben, und uns nicht einmal ganz klar ist, was man eine Lösung nennen könnte, sollen im letzten Abschnitt 4 einige konstruktive Folgerungen gezogen werden.

Da das Problem nicht Produkt einer bestimmten Formalisierung der Entscheidungstheorie ist, werden wir die praktische Überlegung und die Entscheidungstheorie nur in informeller oder halbformaler Weise behan-

² In der klassischen Theorie von Savage (1954) ist die grundlegende Nutzenfunktion für Konsequenzen erklärt, und Handlungen erhalten erwartete Nutzenwerte. In der Standardtheorie der Entscheidungsbäume (vgl. etwa Raiffa, 1968, Kap. 2) ordnet die grundlegende Nutzenfunktion entweder jedem Knoten oder jedem Ast eines Entscheidungsbaumes einen Wert zu. In einem von den Endpunkten zum Ursprung des Baumes laufenden rekursiven Verfahren wird dann jedem Knoten im Hinblick auf dem bei ihm beginnenden Teilbaum ein erwarteter Nutzenwert zugewiesen. In der wieder etwas anderen Theorie von Jeffrey (1965) werden schließlich vollständigen konsistenten Geschichten oder möglichen Welten grundlegende nicht-probabilistische Nutzenwerte zugeordnet, und Propositionen, die aus mehr als einer möglichen Welt bestehen, erhalten erwartete Nutzenwerte.

deln. Doch lassen sich unsere Ausführungen formal leicht mit Hilfe von Entscheidungsbäumen oder Entscheidungsflußdiagrammen nachvollziehen, wie sie in Raiffa (1968) exemplarisch dargestellt sind. Seine Theorie kann daher nötigenfalls als formale Leitlinie dienen; und auf sie beziehen sich unsere Aussagen zur formalen Entscheidungstheorie, auch wenn unsere Intentionen natürlich allgemeiner sind.

Wie sich zeigen wird, ist das hier behandelte Problem nichts anderes als eine Variante von Joseph Butlers Kritik am Hedonismus.³ So gesehen laufen unsere Ausführungen darauf hinaus, daß die modernen Theorien der praktischen Überlegung Butlers Kritik nicht ernsthaft aufgenommen haben.

1. Eine kausale Voraussetzung der Entscheidungstheorie

Unserer These (A) zufolge enthält die Entscheidungstheorie eine versteckte Voraussetzung über die kausale Struktur von Entscheidungssituationen.⁴ Betrachten wir, um diese Voraussetzung aufzudecken, wie wir in Entscheidungssituationen praktisch überlegen. Wir tun dies, um eine Handlung zu finden, die unsere Ziele optimal befördert. Dazu sind einerseits die Ziele, d. h. die intrinsisch gewünschten Sachverhalte, und andererseits die möglichen Handlungen aufzulisten. Jedem Ziel kommt ein Nutzenwert zu, und jedem Ziel und jeder Handlung eine bedingte Wahrscheinlichkeit, nämlich die Wahrscheinlichkeit dafür, daß sich das Ziel realisiert unter der Bedingung, daß die Handlung ausgeführt wurde. Daraus resultiert für jede Handlung ein erwarteter Nutzen. Und so läßt sich schließlich eine Handlung mit maximalem erwarteten Nutzen bestimmen, die also unsere Ziele optimal befördert.

Der erwartete Nutzen von Handlungen scheint hier in einem einzigen Schritt berechnet zu werden. In der Regel ist das aber nicht der Fall; wir gehen normalerweise in unseren Überlegungen von einem sehr viel reicheren kausalen Bild der Welt aus, welches die vielfältigen Wege berücksichtigt, auf denen unsere Handlungen unsere Ziele beeinflussen. Dies

³ Vgl. Butler (1729), Preface, §§ 29–31 und Sermon XI, §§ 1–10. Vgl. auch H. Sidgwick (1874), der sich mit dieser Kritik im Buch I, Kap. IV und Buch II, Kap. III, § 5 auseinandersetzt.

⁴ Es gibt zumindest noch eine andere Voraussetzung dieser Art. Aufgrund dessen, was als ‚Newcombs Problem‘ in die Literatur eingegangen ist, ist die Entscheidungstheorie in zwei Versionen zerfallen, die sich nur unter speziellen Voraussetzungen wieder decken; vgl. die Aufsätze in Campbell/Sowden (1985).

schlägt sich in einer vollständigeren Beschreibung der praktischen Überlegung nieder:

Danach ist es ihr Zweck, wie zuvor den Gesamtnutzen jeder möglichen Handlung und, wie gleich zu sehen sein wird, zudem jedes in Betracht gezogenen Sachverhalts zu bestimmen. Ganz natürlich ist dabei die Annahme, daß der Gesamtnutzen $U'(A)$ eines Sachverhalts A einfach die Summe seines intrinsischen Nutzens $U(A)$ und seines extrinsischen Nutzens $U^e(A)$ ist;⁵ diese Annahme kann freilich nicht als Definition gelten, da sie, wie sich noch herausstellen wird, falsch ist. Man beachte, daß ein Sachverhalt sowohl intrinsischen wie auch extrinsischen Nutzen haben kann: Für manche Leute ist Geld nicht bloß ein Mittel zu einem angenehmen Leben, sondern besitzt schon an und für sich einen Zauber; Autofahren bringt einen schnell, aber unter Lebensgefahr dahin, wo man hin will, und daneben macht es, gelegentlich, als solches Spaß.

Den Ausgangspunkt der praktischen Überlegung bilden demnach die intrinsischen Nutzenwerte der in Betracht gezogenen Sachverhalte (von denen viele gleich Null sein werden). Dann werden auf rekursive Weise der extrinsische und der Gesamtnutzen jedes Sachverhalts bestimmt. Der extrinsische Nutzen eines Sachverhalts A ist ein erwarteter Nutzen; er ist das gewichtete Mittel der Gesamtnutzen der von A direkt kausal beeinflussten Sachverhalte, wobei die Gewichte in den durch A bedingten subjektiven Wahrscheinlichkeiten dieser Sachverhalte bestehen. Damit kann die Handelnde ihr kausales Bild von seinen Endpunkten, d. h. von der Zukunft her aufrollen und so Schritt für Schritt den extrinsischen Nutzen und mit Hilfe der genannten Annahme auch den Gesamtnutzen eines jeden Sachverhalts ermitteln. Zuletzt erreicht sie die Anfangspunkte ihres kausalen Bildes, d. h. die ihr offenstehenden Handlungen und kann so schließlich eine Handlung mit maximalem Gesamtnutzen bestimmen.

Wie auch immer dieses Bild der praktischen Überlegung im einzelnen formalisiert wird,⁶ zentral ist die in ihm enthaltene Konzeption von Wün-

⁵ Die Nutzenfunktion bezieht sich natürlich immer auf eine bestimmte Person zu einer bestimmten Zeit, auch wenn das nicht eigens vermerkt ist.

⁶ Diese Skizze stimmt mit den üblichen Versionen der Entscheidungstheorie gut überein. Das einschrittige Verfahren entspricht der anfänglichen Darlegung der Entscheidungstheorie von Savage (1954) im Kap. 2, wo Entscheidungssituationen mit Weltzuständen, Handlungen und Konsequenzen erfaßt werden. Das mehrschrittige Verfahren wird mit Hilfe von Entscheidungsbaumen formalisiert – vgl. Raiffa (1968), Kap. 2; es ist übrigens auch schon in der Theorie der sogenannten kleinen Welten von Savage (1954), Abschn. 5.5, enthalten – vgl. dazu auch Spohn (1978), Abschn. 2.3 und 3.6. Unsere Ausführungen sind auf diese Standardformalisierung abgestimmt. Schließlich tritt der kausale Charakter der praktischen Überlegung in allen Versionen der kausalen Entscheidungstheorie klar zutage, die ja als eine Ausarbeitung von Savages Theorie intendiert sind; vgl. etwa Skyrms (1984), Kap. 4.

schen oder Nutzenwerten sowie der zwischen ihnen bestehenden Ableitungsbeziehung: Es gibt nur intrinsischen und extrinsischen Nutzen und Summen davon; die intrinsischen Nutzenwerte sind nicht abgeleitet, die extrinsischen sind abgeleitet, und die Ableitung schreitet von den vermuteten kausalen Folgen und ihren Wahrscheinlichkeiten und Gesamtnutzenwerten zu den vermuteten Ursachen und deren extrinsischen Nutzenwerten fort.

Daß wir von Ableitung reden, mag den Eindruck erwecken, als faßten wir die Entscheidungstheorie als eine syntaktische Theorie der rationalen Berechnung von Nutzenfunktionen (und Wahrscheinlichkeiten) auf. Dem ist nicht so. Nach unserem Verständnis formuliert die Entscheidungstheorie vielmehr Koexistenzgesetze für Nutzen- und Wahrscheinlichkeitsfunktionen im Rahmen rationaler Überzeugungs- und Wunschkomplexe. Der eher bildliche Ausdruck, daß ein Teil eines solchen Komplexes sich aus anderen Teilen herleitet, soll hier also durchweg nichts anderes bedeuten, als daß dieser Teil gemäß den entscheidungstheoretischen Gesetzen durch die anderen eindeutig bestimmt ist.

Wenn nun die praktische Überlegung wie eben beschrieben funktionieren soll, so müssen einige Bedingungen kausaler Wohlgeordnetheit erfüllt sein. Eine natürliche Annahme besteht darin, daß der in die praktische Überlegung eingehende Komplex von Wünschen und Überzeugungen hinreichende Ursache des tatsächlichen Verhaltens ist. Diese Annahme gilt per definitionem für alle rational Handelnden.⁷ Sie kann insofern als Voraussetzung jeder praktischen Überlegung gelten, als die Überlegung umso witzloser ist, je weniger sie aufs Handeln durchschlägt. Aus der Verletzung dieser Voraussetzung resultieren Anomalien, die unter dem Stichwort ‚Willensschwäche‘ viel diskutiert werden.⁸

Eine andere Voraussetzung — und um sie geht es hier — lautet, daß das Ergebnis eines jeden Überlegungsschrittes nicht schon bei einem früheren Überlegungsschritt vorausgesetzt werden darf; andernfalls geriete die gesamte Überlegung offenbar in einen Zirkel. Das Ergebnis eines Überlegungsschrittes sind extrinsische Nutzen- und Gesamtnutzenwerte. Wie ist es angesichts der Azyklizität der Kausalbeziehung möglich, daß diese bei einem früheren Überlegungsschritt benötigt werden? Das ist

⁷ Die Kausalisten unter den Handlungstheoretikern sind jedenfalls der Auffassung, daß praktische Gründe Handlungen verursachen können und dies im Falle rationalen Handelns auch tun. Vgl. etwa die ersten drei Aufsätze in Davidson (1980) oder Churchland (1970).

⁸ Vgl. u. a. den zweiten Aufsatz in Davidson (1980), Pears (1986) und Kusser (1989), Abschn. 3.1 und 4.2.

offenbar gerade dann der Fall, wenn die Tatsache, daß ein Sachverhalt für ein Subjekt einen bestimmten extrinsischen oder Gesamtnutzen hat, für diejenigen Wirkungen dieses Sachverhalts kausal relevant ist, aus deren Nutzen sich sein extrinsischer und sein Gesamtnutzen herleitet.

Betrachten wir zunächst ein einfaches abstraktes Beispiel dieser, wie es scheint, absonderlichen Konstellation. Nehmen wir dazu an, daß der Gesamtnutzen von B und $\sim B$ ⁹ für ein Subjekt bereits bestimmt und A der einzige Sachverhalt sei, der für B und $\sim B$ kausal relevant ist. Die Ableitung des extrinsischen und des Gesamtnutzens von A liegt dann auf der Hand:

- (1) $U'(A) = U'(B)P(B|A) + U'(\sim B)P(\sim B|A)$,
- (2) $U'(A) = U^i(A) + U^e(A)$.

Nehmen wir nun weiter an, daß B aus der Sicht der Handelnden unter anderem auch davon kausal abhängt, welchen Gesamtnutzen A für sie hat. In diesem Fall wäre die Wahrscheinlichkeit $P(B|A)$ in (1) nicht zu verwenden; vielmehr müßte man in der Herleitung für alle möglichen Werte x von $U'(A)$ die Wahrscheinlichkeiten $P(B|A \ \& \ U'(A) = x)$ heranziehen.¹⁰ Dann wird jedoch die entscheidungstheoretische Herleitung offenkundig zirkulär; denn um den Wert $U'(A)$ entsprechend dem modifizierten (1) herzuleiten, muß man bereits den Wert von $U'(A)$ kennen, der aber, (2) zufolge, davon abhängt, welchen Wert $U^e(A)$ annimmt. Demnach kann also die entscheidungstheoretische Nutzenherleitung nur funktionieren, wenn eine solche fatale kausale Abhängigkeit nicht besteht.

Wir wollen natürlich behaupten, daß diese Abhängigkeit weder unmöglich noch absonderlich, sondern im Gegenteil ganz normal ist. Doch sind zunächst noch einige Bemerkungen auf abstrakter Ebene am Platze:

Erstens geht unser Argument von der Annahme aus, daß kausale Abhängigkeiten sich in der Sicht der Handelnden in stochastischen Abhängigkeiten gemäß ihren subjektiven Wahrscheinlichkeiten spiegeln. Wir halten diese Annahme für korrekt, doch ist hier nicht der Ort, um für diesen schwierigen Punkt zu argumentieren.¹¹

Zweitens beruht das Argument auf der Annahme, daß die Wahrscheinlichkeit $P(B|A \ \& \ U'(A) = x)$ verwendet werden muß, weil $P(B|A)$ nicht zur Verfügung steht. Das scheint falsch zu sein. Warum kann man

⁹ Die Tilde steht für die Negation von Sachverhalten.

¹⁰ „ $U'(A) = x$ “ steht für den Sachverhalt, daß der Gesamtnutzen von A x ist.

¹¹ Zu unserer Ansicht über den Zusammenhang zwischen Kausalität und Wahrscheinlichkeit vgl. Spohn (1978), Abschn. 3.2+3 und 5.1, sowie (1980) und (1990).

nicht $P(B|A)$ durch die Formel von der totalen Wahrscheinlichkeit berechnen, indem man die Summe (oder das Integral) aller Produkte $P(B|A \& U'(A) = x) \cdot P(U'(A) = x|A)$ bildet? Das würde allerdings voraussetzen, daß die Handelnde für den eigenen Gesamtnutzen von A eine Wahrscheinlichkeitsverteilung hat (darauf läuft $P(U'(A) = x|A)$ ja hinaus); und das mutet seltsam an. Wichtiger ist jedoch, daß diese Annahme sich selbst aufhebt; sie würde erlauben, zuerst $P(B|A)$ und dann, mit Hilfe von (1) und (2), $U'(A)$ zu berechnen und so $U'(A)$ mit Sicherheit zu ermitteln. Natürlich spiegelt sich in dieser Selbstaufhebung nichts anderes als die Zirkularität der ins Auge gefaßten Situation.

Diese Sachlage tritt nur dann nicht auf, wenn man sich von vornherein eines selbst-verifizierenden Wertes von $U'(A)$ sicher ist, eines Wertes also, der, wenn man in den Berechnungen von ihm ausgeht, hinterher gerade wieder herauskommt. Das ist die Lösung, die wir im Abschnitt 4 vorschlagen und erläutern werden. Man beachte jedoch, daß dann von einer Ableitung nicht mehr die Rede sein kann, da es nicht nur *einen* sich selbst verifizierenden Wert von $U'(A)$ zu geben braucht. Und selbst wenn es genau einen solchen Wert gäbe, so ist das Verfahren zur Bestimmung des Wertes von $U'(A)$ doch ein ganz anderes als in dem Fall, wo die fatale kausale Abhängigkeit nicht besteht. Wie drastisch diese Lösung das übliche entscheidungstheoretische Bild verändert, läßt sich in vollen Umfang erst ermitteln, wenn klar geworden ist, wie diese zirkuläre Situation eintreten kann (s. Abschnitt 2) und wieso man im Rahmen des üblichen Bildes diesen Zirkel nicht los werden kann (s. Abschnitt 3).

Drittens enthalten (1) und (2) die implizite Annahme, daß der intrinsische Nutzen von A von all den intrinsischen Nutzen unabhängig ist, die in den extrinsischen Nutzen von A eingehen. Diese Annahme ist jedoch nicht wesentlich. Will man Nutzenabhängigkeiten zulassen, so muß man für den oben erläuterten Überlegungsschritt den Gesamtnutzen von $A \& B$ und den von $A \& \sim B$ als bereits gegeben ansetzen; der Gesamtnutzen dieser Sachverhalte ist ja entweder identisch mit ihren intrinsischen Nutzen oder in den vorherigen Überlegungsschritten bereits ermittelt worden. Auf dieser Grundlage erhält man dann den Gesamtnutzen von A allein entsprechend der Formel:

$$(1') \quad U'(A) = U'(A \& B)P(B|A) + U'(A \& \sim B)P(\sim B|A).$$

Nehmen wir aber wiederum an, B sei von dem Gesamtnutzen, den die Handelnde A zuschreibt, kausal abhängig. Dann kommt man mit (1') aus genau den gleichen Gründen in genau die gleichen Schwierigkeiten wie mit (1) und (2). Man verliert also nichts und gewinnt aber an Durchsich-

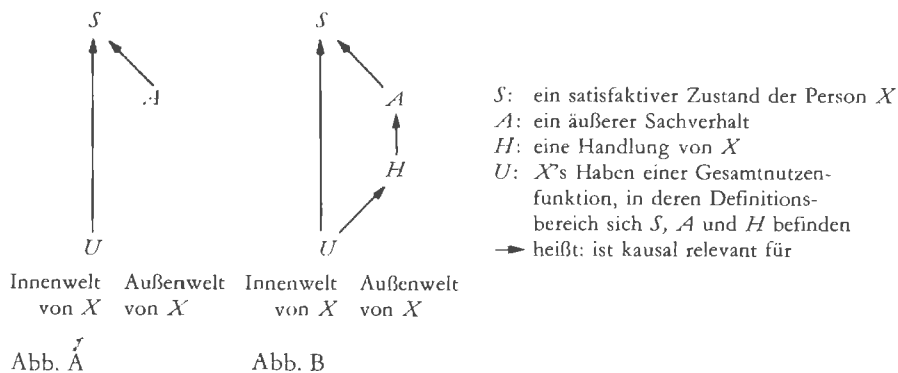
Zustandes; und die die äußere Situation betreffenden Wünsche sind oft relevanter Teil dieses vorhergehenden inneren Zustands.

Es ist nicht schwer, hierfür Beispiele zu finden. Zwei Männer schauen sich im Fernsehen das Lokald Derby an. Die Kickers haben gerade ein Tor geschossen. Der eine ist begeistert, der andere gereizt. Woher kommt dieser Unterschied? Der eine ist eben Kickers-Fan, der andere VfB-Anhänger; für den einen hat das Tor positiven Nutzen, für den anderen negativen, da jeder möchte, daß seine Mannschaft gewinnt.

Ein anderes Beispiel: Vor zwei Jahren war eine Frau schwanger, und sie war ganz unglücklich, nun ist sie wieder schwanger und freut sich sehr. Warum? Seinerzeit wollte sie kein Kind, jetzt hat sie sich ein Kind gewünscht. Das ist eine kausale Erklärung. Zunächst war da der Wunsch nach einem Kind, dann wurde sie schwanger, und nun ist sie glücklich; hätte sie diesen Wunsch, wie vor zwei Jahren, nicht gehabt, hätte die Schwangerschaft nicht diese Freude ausgelöst.

Vorausgehender Nutzen und nachfolgende Befriedigung brauchen jedoch nicht parallel zu laufen. Alle Arten von Zusammenhängen sind möglich. Es kann sein, daß ich ein Ziel so wichtig nehme, daß ich es ganz verbissen verfolge und so am Ende die Früchte meiner Anstrengungen nicht mehr genießen kann. Der verzweifelte Wunsch nach etwas, was man entbehrt, kann sich in zwei Richtungen auswirken: Man kann nicht genug davon bekommen und ist permanent frustriert, oder man freut sich demütig schon an kleinen Mengen. Ja, es gibt sogar Dinge oder Situationen, die nur dann befriedigen, wenn man sie nicht herbeizuführen versucht oder nicht einmal wünscht.¹⁶

All diese Beispiele weisen die im vorhergehenden Abschnitt als problematisch erwiesene Kausalstruktur auf:



¹⁶ Vgl. Elster (1983), der im Abschn. II.2 paradoxe Situationen dieser Art diskutiert.

Die Abbildung A zeigt das Grundgerüst dieser Struktur; es ist im Fußballbeispiel realisiert, wobei A die Tatsache ist, daß die Kickers gerade ein Tor geschossen haben und S für die Begeisterung des einen bzw. für die Gereiztheit des anderen steht. Eine Kausalkette läuft von der äußeren Tatsache A (via Fernsehen usw.) zum inneren Zustand S . Eine zweite Kausalkette läuft von der Nutzenfunktion U zu S , ohne die innere Welt der Person X zu verlassen. Diese im einzelnen nicht ausgeführte Kette, der U als ein Glied angehört, bildet die psychische Grundlage für X 's Reaktion auf die äußeren Ereignisse. Die Abbildung A zeigt weiters keine Handlungen; aber sie demonstriert schon den springenden Punkt, nämlich daß eine zweite Kausalkette existiert, die den erläuterten Zirkel auslöst und die Ableitung des Gesamtnutzens von A aus dem intrinsischen Nutzen von S vereitelt.

Die Abbildung B berücksichtigt nun ausdrücklich auch Handlungen; sie paßt auf das Schwangerschaftsbeispiel, wobei A für die Schwangerschaft der Frau steht, S für ihre Freude und H etwa für die Handlung, Verhütungsmittel abzusetzen. Hier laufen nun von U zu S tatsächlich zwei Kausalketten, die erste über H und A , die zweite wie in der Abbildung A. Wie bereits erläutert, stellt die erste kein Problem dar, weil der kausale Einfluß von U auf S in dieser Kette von H abgeschirmt wird; daher kann hier der extrinsische Nutzen etwa von H aus dem Gesamtnutzen von A (und $\sim A$) abgeleitet werden. Dagegen ist es auch hier wieder die zweite Kausalkette – die nicht durch die äußere Welt abgeschirmte – die dafür verantwortlich ist, daß der Gesamtnutzen von A , und folglich auch der von H , nicht aus dem Nutzen von S abgeleitet werden kann.

Der intuitive Kern von alledem ist, daß die praktische Überlegung nicht von der Stelle kommt, wenn sie einzig vom Wunsch nach Glück ausgeht. Wenn man sich z. B. für einen Beruf zu entscheiden hat, so kommt man zu keinem Ergebnis, wenn man nur danach fragt, wie befriedigend wohl die verschiedenen Optionen wären. Der Grund dafür liegt nicht darin, daß die Entscheidung so komplex wäre, sondern darin, daß die Optionen an sich kein bestimmtes Befriedigungspotential haben. Vielmehr muß man bereits zumindest eine grobe Vorstellung davon haben, wie sehr man die jeweiligen Berufe ausüben will; erst dann kann man sich sinnvoll fragen, wie sehr sie jeweils befriedigen würden. Das war der wesentliche Punkt an Joseph Butlers vor über 250 Jahren geübten Kritik am Hedonismus:

„The very idea of an interested pursuit necessarily presupposes particular passions or appetites, since the very idea of interest or happiness consists in this, that an appetite or affection enjoys its object. It is not because we love ourselves that we find delight in such and such objects, but because we have particular

wenn $U^e(A)$ nicht existiert, ebenfalls nicht existieren kann. Demnach, so dieser Einwand, ist es völlig mysteriös, welcher Nutzen diesen Zirkel erzeugen soll.

Die formale Entgegnung darauf lautet, daß im Abschnitt 1 in Vorwegnahme dieses Einwands betont wurde, daß es nur eine Annahme und keine Definition ist, wenn $U^k(A)$ gleich der Summe von $U^i(A)$ und $U^e(A)$ gesetzt wird. Läßt man diese Annahme fallen, wie es der Zirkel erzwingt, fällt der letzte Schritt des Einwands weg, so daß der Gleichsetzung von $U^k(A)$ mit $U^i(A)$ nichts im Wege steht.

Aus unseren Beispielen ergibt sich aber auch eine intuitive Entgegnung. Die Freude der Frau hängt davon ab, wie sehr sie sich das Baby wünscht; meine berufliche Befriedigung ist dadurch bedingt, wie sehr ich den Beruf, den ich habe, auszuüben wünsche. In diesen Fällen ist der die Befriedigung beeinflussende Wunsch nicht (oder nicht nur) ein intrinsischer. Im Falle der Frau ist es sozusagen ihr Gesamtwunsch, der sich aus vielen Quellen speist: aus dem intrinsischen Wunsch nach einem Kind, aus dem negativen Nutzen der Kosten, die das Kind verursacht; aus dem Wunsch, ihren Mann oder sogar ihre Eltern zufriedenzustellen, etc. und natürlich aus ihrer Freude an dem Kind. Im Falle des Berufes ist es sogar leicht vorstellbar, daß er überhaupt nicht intrinsisch gewünscht wird, sondern nur als Mittel zur Befriedigung, zum Geldverdienen, etc.

Aus diesem Einwand ergibt sich also nicht, daß unsere intuitive Beschreibung des Zirkels falsch ist; vielmehr folgt, daß mit der üblichen Einteilung des Nutzens in intrinsischen, extrinsischen und Gesamtnutzen etwas nicht stimmt. Die Beispiele weisen auf eine Art von Gesamtnutzen hin, der unter den Bedingungen des Zirkels weder intrinsisch noch entscheidungstheoretisch ableitbar ist. Zu suchen ist mithin eine adäquatere Klassifikation des Nutzens.

(IV) Man könnte immer noch den Eindruck haben, daß sich unser Problem einem unklaren Nutzenbegriff verdankt. Wir haben uns nämlich massiv auf die kausale Trennbarkeit von Nutzen und Befriedigung gestützt; und diese mag anfechtbar erscheinen. Tatsächlich wird beides vielfach gleichgesetzt. Diese Gleichsetzung wollen wir nun nach zwei Richtungen hin diskutieren, wobei wir insgesamt auf folgendes hinauswollen: Wenn man (a) Nutzen auf Befriedigung reduziert, wird man dem Nutzenbegriff nicht gerecht; reduziert man umgekehrt (b) Befriedigung auf Nutzen, so wird man dem Phänomen der Befriedigung nicht gerecht. Schließlich gibt es (c) eine Idee zur Überbrückung der Differenz von Nutzen und Befriedigung, die uns attraktiv erscheint; doch bringt sie, wie sich zeigen wird,

keine Lösung des Überlegungszirkels. Betrachten wir (a), (b) und (c) näher.

(IVa) Der Handelnde ist nach der Entscheidungstheorie ein höchst eigenartiges Wesen. Er ist dauernd damit beschäftigt, seinen Nutzen zu berechnen und seinen erwarteten Nutzen zu maximieren – und damit hat es sich. Aber hat er jemals etwas von all der Mühsal? Was ist am Ende seine ‚Auszahlung‘? In den Lehrbuchbeispielen, wo Geld vorläufig die Rolle des Nutzens übernimmt, ist die Antwort einfach: Er bekommt eben Geld. Aber in Wirklichkeit lassen sich Geld und Nutzen nicht so einfach gleichsetzen. Was bekommt er dann? Die natürliche Antwort darauf ist, daß er dann eben den *Nutzen* bekommt; schließlich liegt kein Sinn darin, den Erwartungswert einer Quantität zu maximieren, von der man nichts bekommt. So weit ist der Nutzen allerdings nichts weiter als ein Name für das, was der Handelnde am Ende bekommt. Aber nun wissen wir doch, worin seine ‚Auszahlung‘ besteht; sie besteht darin, wie er sich fühlt und wie es ihm geht, d. h. in unserer Terminologie, in seinem satisfaktiven Zustand, gemessen auf irgendeiner Skala, wie sie schon im alten Lustkalkül konzipiert war. Das also ist sein Nutzen.

So sieht kurz die Überlegung aus, die Nutzen zwanglos auf Befriedigung reduziert; viele haben sich von ihr verführen lassen.¹⁸ Doch sie ist noch unzureichend. Denn so weit besagt sie bloß, daß die zukünftige Auszahlung des Handelnden in seinen zukünftigen satisfaktiven Zuständen bestehen wird. Wie aber ist die gegenwärtige Nutzenfunktion zu verstehen? Es scheint nur eine Möglichkeit zu geben: Der gegenwärtige Nutzen eines möglichen zukünftigen satisfaktiven Zustands ist einfach der Grad der Befriedigung, die mit ihm eintreten würde, und der gegenwärtige Nutzen anderer Sachverhalte besteht in ihrer erwarteten Befriedigung, d. h. er ist ein Erwartungswert, der ausdrückt, für wie befriedigungsfördernd diese Sachverhalte gehalten werden.

Diese Konzeption ist jedoch immer noch unzureichend. Denn das von ihr gezeichnete Bild praktischen Überlegens und rationalen Handelns enthält nur zukünftige Befriedigung und gegenwärtige Überzeugungen darüber; die Wünsche oder Ziele des Handelnden kommen in ihm nicht vor. Es scheint uns klar, daß ein solches Bild grundlegend verzerrt sein muß.

¹⁸ Diese Reduktion ist für den klassischen Utilitarismus charakteristisch, doch findet sie sich auch in der modernen Wohlfahrtstheorie. Vgl. Bentham (1970), Kap. I, Mill (1976), Kap. 2, Sidgwick (1874), Buch II, und Pigou (1920). Vgl. auch Sen (1980/81), Teil 2, und Griffin (1986), Kap. 1.

Man könnte es nachzubessern versuchen, indem man die Wünsche des Handelnden innerhalb dieses Bildes definiert. Das erscheint jedoch von vornherein wenig aussichtsreich, da Wünsche gemeinhin als eine grundlegende Art propositionaler Einstellungen gelten, die sich nicht auf andere propositionale Einstellungen reduzieren lassen.¹⁹ Außerdem könnte man sie wohl nur dadurch ins Bild aufnehmen, daß man den bereits definierten gegenwärtigen Nutzen des Handelnden zugleich als Repräsentation seiner mehr oder weniger starken Wünsche auffaßt. Dieses Vorgehen würde jedoch den Hedonismus in eine Wahrheit a priori verwandeln; denn es postuliert, daß sich die Wünsche des Handelnden intrinsisch ausschließlich auf satisfaktive Zustände richten. Außerdem ergibt sich ein bekanntes Problem des Hedonismus: Um den intuitiven Wunschbegriff mit diesem Postulat in Einklang zu bringen, muß jedweder Wunsch mit einer Befriedigungserfahrung und jedwede Präferenz mit einer Differenz in der (erwarteten) Erfahrung von Befriedigung in Zusammenhang gebracht werden. All das erscheint nicht minder verzerrt.

Der Punkt ist mithin schlicht der, daß gegenwärtige Wünsche nicht mit gegenwärtigen Erwartungen über zukünftige satisfaktive Zustände gleichgesetzt werden können und daß der gegenwärtige Nutzen eben nicht die letzteren, sondern die ersteren repräsentieren muß, wenn er im praktischen Überlegen und im rationalen Handeln die ihm von der Entscheidungstheorie zugeordnete Rolle spielen soll.²⁰

Schließlich ist darauf hinzuweisen, daß die zur Diskussion stehende Konzeption des Nutzens den Überlegungszirkel nicht zum Verschwinden bringt, sondern bloß transformiert. Denn unsere These (B) besagte dann, daß die gegenwärtige Nutzenerwartung des Handelnden für seine zukünftige Befriedigung kausal relevant ist; und dafür wären genau dieselben Beispiele einschlägig. Einfacher und nicht probabilistisch ausgedrückt bedeutet das, daß die gegenwärtige Überzeugung darüber, daß dieser Sachverhalt zu jener Befriedigung führen wird, kausal dafür relevant ist, ob das tatsächlich der Fall ist. Diese Überzeugung würde sich also selbst bewahrheiten oder jedenfalls ihre eigene Wahrheit kausal affizieren. Und daraus entstünde dann ein analoger Überlegungszirkel: Um zu wissen, welche zukünftige Befriedigung man erwarten soll, sollte man bereits wissen, welche Erwartung man diesbezüglich hat. Diese Konzeption des

¹⁹ Vgl. auch Lewis (1988), der ein allgemeines Argument dafür bringt, daß sich Wünsche nicht auf Überzeugungen reduzieren lassen.

²⁰ Dies wird in Kusser (1989), Abschn. 3.3, ausführlicher begründet.

Nutzens ist also, um ein Fazit zu ziehen, nicht nur fehlgeleitet, sie löst auch nichts – Gründe genug, um von ihr Abstand nehmen.²¹

(IVb) Tatsächlich ist die Geschichte des Nutzenbegriffs in diesem Jahrhundert durch diesen Abstand geprägt, wie er sich vor allem in der Theorie der ‚revealed preferences‘ zeigt.²² Die meisten Theoretiker und insbesondere die Ökonomen begreifen heute Nutzenfunktionen als Repräsentation der unterschiedlich starken Wünsche, wie sie sich in den Präferenzen und den Handlungen rationaler Personen manifestieren.²³ Dann stellt sich jedoch erneut die Frage nach dem Verhältnis von Nutzen und Befriedigung.

Da gibt es nun die zähe Neigung, die soeben diskutierte Antwort umzudrehen und zu sagen, daß Befriedigung eben das ist, was sich einstellt, wenn Wünsche in Erfüllung gehen – womit der Befriedigungsbegriff zu einem rein formalen wird.²⁴ Das klingt absurd, und vermutlich hat es deshalb niemand so deutlich gesagt. Doch hat die Neigung, Befriedigung (oder Lust oder Glück) so zu konzipieren, eine reiche Geschichte. Kant etwa sagt:

„Glückseligkeit ist der Zustand eines vernünftigen Wesens in der Welt, dem es im Ganzen seiner Existenz alles nach Wunsch und Willen geht.“²⁵

Und Rawls schreibt:

²¹ Es ist anzumerken, daß der von G. S. Becker und K. Lancaster u. a. entwickelte und vor allem unter Ökonomen prominente Ansatz, auch Haushalten Produktionsfunktionen zuzuordnen (vgl. Michael/Becker, 1973), nicht minder von unserer Schwierigkeit betroffen ist. Im Rahmen dieses Ansatzes wird der Konsument als Produzent von Nutzen oder Lust oder Befriedigung angesehen, was die Möglichkeit eröffnet, die Produktionstheorie auf den Konsumenten anzuwenden. Unser Zirkel impliziert, daß die Produktionsfunktion des Konsumenten für ein gewisses Grundgut (d. h. eine bestimmte Art von Befriedigung) in der Regel ein zusätzliches Argument enthalten muß, nämlich die Stärke seines Wunsches nach denjenigen Gütern, aus denen er das Grundgut produziert; und nach dem gerade diskutierten verzerrten Bild wäre ein anderes Argument zusätzlich zu berücksichtigen, nämlich die Überzeugungen darüber, wieviel vom Grundgut aus welchen Mengen anderer Güter produziert werden kann. So oder so ergibt sich daraus eine Schwierigkeit oder Zirkularität, die in der herkömmlichen Produktionstheorie nicht entsteht und nicht behandelt wird.

²² Die Gründungsväter dieser Theorie sind v. Neumann/Morgenstern (1944), Abschn. 1.3, und Samuelson (1947), Kap. IV.

²³ Positivistischer Gesinnung halten in der Tat nur Handlungen und Präferenzen für real und reden so, als ob Wunschstärken und Nutzenwerte nur Als-ob-Konstruktionen wären. Ein solcher Operationalismus gilt aber in der gegenwärtigen Wissenschaftstheorie als überholt; vgl. auch Sen (1973), Abschn. 2.

²⁴ Dem entspricht ein rein formaler Begriff der Auszahlung als eines Nutzens, den man immer dann schon hat, wenn ein Wunsch in Erfüllung geht. Vgl. Griffin (1986), Kap. 1, und Sen (1980/81), Teil 4.

²⁵ Kant (1788), S. 224.

„... we are happy when our rational plans are going well, our more important aims being fulfilled, and we are with reason quite sure that our good fortune will continue.“²⁶

Gleichwohl sollte man dieser Neigung nicht nachgeben. Zu häufig war man nicht befriedigt, obschon man bekommen hat, was man wollte, und zu lebhaft und eigenständig ist die Phänomenologie der satisfaktiven Zustände, als daß die begriffliche Verschmelzung von Wunscherfüllung und Befriedigung durchführbar wäre.

Natürlich gibt es Gründe für diese Neigung. Einer ist sicherlich, daß es zumindest die Regel ist, daß Wunscherfüllung befriedigt; doch es gibt Ausnahmen. Ein besserer Grund ist, daß wir nicht umhinkönnen, jeweils anzunehmen, daß die Erfüllung unserer Wünsche uns befriedigen wird. Sind wir der Überzeugung, daß die Erfüllung eines Wunsches uns nicht befriedigen wird, und halten wir gleichwohl an diesem Wunsch fest, so haben wir entweder überwiegende, z. B. moralische Gründe dafür, oder wir sind einfach töricht. Die Situation ähnelt Moores Paradox: Wir müssen jeweils glauben, daß das, was wir glauben, wahr ist; doch wissen es andere vielleicht besser und wir später womöglich auch. Genauso müssen wir jeweils glauben, daß das, was wir wünschen, uns befriedigt; aber andere wissen es möglicherweise besser, und wir merken es vielleicht auch noch. Diese Gründe mögen die diskutierte Neigung genährt haben; doch ist sie dadurch nicht gerechtfertigt.

(IVc) Hat man erst einmal eingeräumt, daß der Zusammenhang zwischen Befriedigung und Wünschen oder Nutzen kontingent ist, so kommt man um den Zirkel nicht mehr herum. Zwar wird immer wieder gesagt, daß unsere wirklichen oder wahren Wünsche jene seien, deren Verwirklichung befriedigt, und daß wir im Idealfall diese Wünsche hätten.²⁷ Wir teilen diese Ansicht und werden unten darauf zurückkommen. Der Überlegungszirkel ist davon allerdings nicht betroffen, da er auf dem tatsächlichen Zusammenhang zwischen Wünschen und Befriedigung beruht und nicht durch die Betrachtung des Idealfalls verschwindet.

(V) Gleichwohl mögen unsere Zirkelthesen immer noch ein ungutes Gefühl zurückgelassen haben, das vielleicht im folgenden Einwand zum

²⁶ Rawls (1971), S. 548.

²⁷ Vgl. etwa Russell (1921), Kap. III, insbes. S. 72 ff, Gosling (1969), Kap. 7, Tugendhat (1984), S. 45 f, und Tversky/Kahneman (1986), S. 138. — Von einer rationalen Person ist nicht zu verlangen, daß sie wahre Wünsche hat, genauso wenig, wie sie wahre Überzeugungen haben muß. Zur Rationalität gehört nicht, Wissen zu haben, sondern Wissen zu suchen (wobei natürlich das erste häufig Resultat des zweiten ist). Ebenso gehört es sicherlich zur Rationalität, nach wahren Wünschen zu streben. Vielleicht ist dieser Sinn von Rationalität in den obigen Zitaten von Kant und Rawls intendiert.

Ausdruck kommt: Nach der These (B) soll sehr häufig etwas eine Person befriedigen, *weil* sie es gewünscht hat, und darin soll dann irgendwie ein Problem liegen. Fragt man aber die Person, warum sie die Sache wünscht, so ist die natürliche Antwort: „*Weil* es mich befriedigt!“ Und an dieser Antwort ist nichts auszusetzen – außer, daß offenbar nicht beides, diese Antwort und die These (B), richtig sein kann.

Dieser Einwand kann einen leicht verwirren. Doch besteht hier kein Widerspruch; das „*weil*“ wird dabei in zwei verschiedenen Bedeutungen verwandt. Sagt man, etwas befriedige eine Person, weil sie es wünsche, so redet man über eine der Ursachen ihrer Befriedigung; sagt sie jedoch, sie wünsche es, weil es sie befriedige, so spricht sie über einen ihrer praktischen Gründe für ihren Wunsch.

Ganz so leicht ist die Sache aber nicht abgetan; es gibt hier noch zwei Probleme:

(Va) Wir haben gerade zugegeben, daß die Überzeugung, daß etwas befriedigen wird, ein praktischer Grund dafür ist, es zu wünschen. Was aber ist ein praktischer Grund? Ist *A* nicht genau dann ein praktischer Grund für *B*, wenn *A* in einer zu *B* führenden praktischen Überlegung, d. h. also in einer zu *B* führenden entscheidungstheoretischen Ableitung wesentlich verwandt wird? Doch enthält der letzte Satz zwei verschiedene, durch das „d. h.“ zu unrecht gleichgesetzte Erläuterungen. Die erste ist vage und richtig, die zweite präzise und falsch. Man wird also nicht umhinkommen, zwei Arten praktischer Gründe anzuerkennen: die in der Entscheidungstheorie explizierten und noch andere; und zu dieser anderen Art gehört offenbar die Überzeugung, daß etwas befriedigen wird. Es bleibt die Aufgabe, diese Art von praktischen Gründen genauer zu erklären.

(Vb) Die Kausalisten unter den Handlungstheoretikern, denen wir uns, wie klar sein dürfte, angeschlossen haben, behaupten, daß sich bei einer rationalen Person das, was aus ihrer Innenperspektive ein praktischer Grund ist, aus der Außenperspektive als eine Ursache darstellt. Da wir einräumen, daß die Überzeugung, daß etwas befriedigen wird, für eine rationale Person ein Grund ist, es zu wünschen, müssen wir damit auch zugeben, daß diese Überzeugung ihren Wunsch verursacht. Da wir außerdem behauptet haben, daß ihr Wunsch für ihre Befriedigung kausal relevant ist, folgt daraus (mit der Transitivität kausaler Relevanz), daß die Überzeugung, daß etwas befriedigen wird, für ihre eigene Wahrheit kausal relevant ist. Soll man diese Folgerung wirklich akzeptieren?²⁸

²⁸ Diese Problemliste ist natürlich in keiner Weise erschöpfend. Eine wichtige, wenn auch

4. *Wie man dem Zirkel gerecht werden kann*

So weit scheint der Zirkel ungebrochen zu sein, doch hat die Diskussion etliche offene Fragen aufgeworfen. Ist der Zirkel wichtig, so fordert er eine ganze Theorie. Die haben wir nicht, aber einige konstruktive Beobachtungen und positive Schlußfolgerungen möchten wir schließlich noch anbieten.

Bereits im Abschnitt 1 haben wir auf die Rudimente einer formalen Lösung des Problems hingewiesen. Betrachten wir, um sie etwas zu verdeutlichen, noch einmal die Formeln (1) und (2):

- (1) $U^e(A) = U^e(B)P(B|A) + U^e(\sim B)P(\sim B|A)$,
 (2) $U^i(A) = U^i(A) + U^e(A)$.

In ihnen ist sozusagen ein reguläres Verhältnis von A und B unterstellt. Im zirkulären Fall, der, wie gesagt, nur dann eintritt, wenn B für einen zukünftigen satisfaktiven Zustand steht, sind die kausalen Überzeugungen der Handelnden allerdings nicht durch $P(B|.A)$ zu repräsentieren, sondern durch $P(B|.A \ \& \ U^i(A) = x)$ mit einem geeignet gewählten Wert von x . Also sollte man in (1) die erste Wahrscheinlichkeit durch die zweite ersetzen. Wenn nun bekannt wäre, daß $U^i(A)$ den Wert x annimmt, so ergibt sich aus den modifizierten Formeln (1) und (2) ein Wert y für $U^e(A)$. Das heißt, wenn der Wert von $U^i(A)$ x ist, so sollte er y sein, was nur erfüllbar ist, wenn $x = y$ ist. Mit anderen Worten: Nur Lösungen der Gleichung

- (3) $x = U^i(A) + U^e(B|A \ \& \ U^i(A) = x)$
 $+ U^e(\sim B)P(\sim B|A \ \& \ U^i(A) = x)$

ganz andere Variation des Einwands (IV) besagt, daß wir von einer problematischen Trennung in den Gegenständen oder Inhalten von Wünschen ausgehen. In der Tat haben wir implizit angenommen, daß es z. B. über den Wunsch nach einem bestimmten Bordeaux hinaus noch den Wunsch nach dem Geschmack dieses Bordeaux gibt, d. h. allgemein, daß der Inhalt eines Wunsches in zwei Teile zerlegt werden kann, einen externen Teil, der sich auf äußere Ereignisse bezieht, und einen internen, der nur von inneren Erfahrungen handelt. Dagegen könnte man einwenden, daß diese Trennung unzulässig sei, weil der Inhalt eines Wunsches im Prinzip beschreibbar sein müsse und weil satisfaktive Zustände sich oftmals nur unter Bezugnahme auf die äußere Welt beschreiben lassen (z. B. auf einen bestimmten Jahrgang einer bestimmten Weinlage). Und man könnte ferner vermuten, daß wir den Zirkel genau aus dieser Trennung gewinnen, indem wir einen Wunsch in zwei aufspalten und dann aus dem Zusammenhang der beiden Hälften ein Problem konstruieren. Demgegenüber halten wir diese Trennung für legitim; dies zu verweigern, würde uns freilich (zu) weit in die Philosophie des Geistes und der Sprache abführen — d. h. genauer, in die Debatte um den Antidualismus von Burge (1979), wo in Bezug auf Überzeugungen analoge Fragen diskutiert werden.

sind zulässige Werte von $U'(A)$; andernfalls entspricht der Gesamtnutzen von A nicht dem Erwartungswert der Befriedigung, den A in Bezug auf B unter der Annahme dieses Gesamtnutzens hat.

In der Tatsache, daß die Unbekannte x auf der linken und der rechten Seite der Gleichung (3) erscheint, spiegelt sich noch einmal unmittelbar der Überlegungszirkel. Man beachte, daß A nun kein extrinsischer Nutzen mehr zugeschrieben wird; wenn der Gesamtnutzen der Formel (3) gehorcht, so läßt er sich nicht mehr als Summe von intrinsischem und extrinsischem Nutzen verstehen. Die Gleichung (3) hat keine, genau eine oder, was häufig der Fall sein wird, mehrere Lösungen, je nachdem wie die bedingten Wahrscheinlichkeiten in (3) mit x variieren. Wenn das mathematische Problem sauber formalisiert wäre, könnte man darauf wie in der Gleichgewichtstheorie Fixpunkttheoreme anwenden. Wir sind dem aber nicht nachgegangen, aus Gründen, die weiter unten zur Sprache kommen.

Gleichwohl ist der in (3) zum Ausdruck kommende Gleichgewichtsgedanke intuitiv sehr hilfreich; sechs offen gebliebene Fragen lassen sich auf seiner Grundlage informell klären.

Vier Arten von Nutzen: Die Gleichung (3) stellt zwischen dem Nutzen satisfaktiver Zustände und dem äußerer Sachverhalte einen Zusammenhang her, wie wir ihn im Teil (II) des vorigen Abschnitts gesucht haben. Dieser Zusammenhang ergänzt das entscheidungstheoretische Bild praktischer Rationalität. Innerhalb des alten Bildes gibt es nur zwei Arten von Nutzen: intrinsische Nutzenwerte, von denen alle Ableitungen ausgehen, und abgeleitete, d. h. extrinsische und Gesamtnutzenwerte. In (3) hingegen sind die Rollen von Nutzenwerten anders verteilt. Man kann nicht sagen, daß in (3) $U'(A)$ aus $U'(B)$ oder $U'(B)$ abgeleitet wäre.²⁹ Das ist besonders deutlich, wenn (3) mehr als eine Lösung hat. Der intrinsische Nutzen satisfaktiver Zustände, wie B einer ist, dient also in der Regel nicht als Prämisse einer Ableitung. Entsprechend ist der Nutzen vieler äußerer Sachverhalte weder abgeleitet noch im ursprünglichen Sinn intrinsisch. Vielmehr kann man das Verhältnis zwischen $U'(B)$ oder $U'(B)$ und $U'(A)$ in (3) als eines der Kontrolle bezeichnen. Der Nutzen äußerer Sachverhalte wird in dem Sinne von dem Nutzen satisfaktiver Zustände kontrolliert, daß der erstere mit dem letzteren im Gleichgewicht sein muß

²⁹ Wenn B , wie in (3), ein satisfaktiver Zustand ist, so werden $U'(B)$ und $U'(B)$ gewöhnlich zusammenfallen.

und zu verändern ist, sobald ein Ungleichgewicht entstanden ist.³⁰ Nutzenfunktionen können mithin auch einer solchen Gleichgewichtskontrolle unterliegen; das ist die wesentliche Lehre, die aus dem Überlegungszirkel zu ziehen ist.

In diesem Sinn ergeben sich aus (3) zwei neue Arten von Nutzenwerten: kontrollierende und kontrollierte. Damit klärt sich auch die in (III) aufgetretene Verwirrung darüber, welcher Art der für Befriedigung kausal relevante Nutzen $U^k(\mathcal{A})$ von \mathcal{A} ist. Es ist der unabgeleitete Gesamtnutzen von \mathcal{A} – wobei „unabgeleitet“ nun nicht mehr „intrinsisch“ im alten Sinn und „Gesamtnutzen“ nicht „abgeleitet“ bedeutet. Darin liegt nichts Paradoxes, wenn man nur einsieht, daß es diese vier Arten von Nutzen gibt.³¹

Zwei Arten praktischer Gründe: Das eben Gesagte beantwortet auch die Frage in (Va) des vorstehenden Abschnitts. Es gibt tatsächlich zwei Arten praktischer Gründe: jene, die als Prämissen in entscheidungstheoretischen Ableitungen fungieren, und jene, die Gleichgewichte wie die Lösungen von (3) stützen. Der Wunsch nach Befriedigung und der Glaube, daß etwas befriedigen wird, oder, allgemeiner, Nutzen und Wahrscheinlichkeiten wie diejenigen auf der rechten Seite von (3) sind Gründe dieser zweiten Art. Auf die Frage, warum man etwas wünscht, mit „weil es mich befriedigt“ zu antworten, ist also völlig richtig. Falsch wäre nur die Annahme, daß dieser Wunsch damit aus dem Wunsch nach Befriedigung abgeleitet ist.

³⁰ Tatsächlich ist dieses Verhältnis weniger einseitig, als das Wort „Kontrolle“ andeutet. Erlebt man z. B., daß eine gewisse Sache nicht mehr in der früheren Weise befriedigt, so muß man nicht unbedingt den Wunsch nach dieser Sache aufgeben; man kann auch versuchen, sie in neuer Weise schätzen zu lernen, d. h. sich so zu verändern, daß man eine andere Befriedigung aus ihr zieht – indem man z. B. seine Aufmerksamkeit auf anderes lenkt, die alten Arten von Befriedigung geringer schätzt, usw. Man kann also ein Gleichgewicht auch dadurch wiederherstellen, daß man satisfaktive Zustände neu evaluiert bzw. sich dahin bringt, daß die gleichen äußeren Ereignisse andere satisfaktive Zustände hervorrufen. (Letzteres ist in der Tat die angemessenere Beschreibungsweise, da ein satisfaktiver Zustand durch seine Evaluation wesentlich charakterisiert ist und mit einer neuen Evaluation ein anderer wird; eine nicht mehr bloß positiv bewertete Erregung etwa *ist* nicht mehr die alte Erregung). Diese Anpassung der Freuden an die Wünsche (anstatt der Wünsche an die Freuden) ist tatsächlich häufig zu finden; man sollte sie nicht als irrational denunzieren.

³¹ Hierbei haben wir nur die sozusagen reinen Arten von Nutzen erwähnt. Sie können sich auch mischen, und dann wird es wirklich kompliziert. So mag z. B. für die schwangere Frau das Kind, das sie erwartet, alle drei Arten von Nutzen haben: einen intrinsischen, weil sie neues Leben positiv bewertet; aus vielerlei Gründen einen extrinsischen Nutzen, etwa weil es viel Geld kosten wird; und schließlich noch jener von ihrem Wunsch, glücklich zu sein, kontrollierte Nutzen.

In gleicher Weise ist die Behauptung zu qualifizieren, daß das, was aus der Innenperspektive ein praktischer Grund ist, von außen betrachtet eine Ursache ist. Diese Behauptung läßt sich nur für die entscheidungstheoretisch funktionierenden Gründe, für die sie ja auch intendiert war, aufrecht erhalten. Die ekuilibristischen praktischen Gründe verursachen hingegen nur, daß man an einem durch sie gestützten Wunsch festhält, nicht jedoch seinen Erwerb. Auf diese Weise läßt sich die unerfreuliche Schlußfolgerung aus (Vb) vermeiden. Gleichwohl müssen wir uns diesen Punkt noch genauer anschauen.

Der Status der Überzeugung, daß die Verwirklichung eines Wunsches befriedigen wird: Nach dem eben Gesagten begleitet diese Überzeugung den Wunsch eher, als daß sie ihm vorausgeht. In der Tat begleitet sie ihn notwendig; gibt es keine anderweitigen Gründe für einen Wunsch, so wäre es irrational, diesen Wunsch zu haben, ohne der Überzeugung zu sein, daß seine Erfüllung befriedigen wird. Das, oder vielmehr dessen probabilistische Entsprechung, ist der Inhalt von (3); und dies stimmt mit dem überein, was wir in (IVb) festgestellt haben.

Was ist aber mit dem in (Vb) und am Ende von (IVa) aufkeimenden Verdacht, daß diese Überzeugung keine normale empirische, sondern eine sich selbst bewahrheitende Überzeugung ist? Hier gilt es den genauen Inhalt der Überzeugung zu beachten. Er kann gegeben sein durch (a): „die Realisierung des gewünschten Sachverhalts \mathcal{A} wird befriedigen“ oder durch (b): „die Realisierung des Sachverhalts \mathcal{A} wird, sofern er gewünscht wird, befriedigen“. Jeder, der (a) glaubt, wird vermutlich auch (b) glauben; und jeder, der (b) glaubt und sich des Wunsches nach \mathcal{A} bewußt ist, wird sicherlich auch (a) glauben. Und trotzdem besteht zwischen (a) und (b) eine subtile Differenz. Wenn man der kausalen Überzeugung (a) ist und dabei übersieht oder leugnet, daß der Wunsch nach \mathcal{A} für die Befriedigung durch \mathcal{A} kausal relevant ist, dann kann man zwar das bisherige Bild aufrecht erhalten, aber nur um den Preis, daß die Überzeugung (a) ihre eigene Wahrheit kausal beeinflusst. Ist man hingegen der kausalen Überzeugung (b), so muß man zwar in Bezug auf den Nutzen von \mathcal{A} das alte Bild aufgeben, hat aber den Vorteil, daß (b) nun eine normale empirische Überzeugung darstellt. Nur die zweite Alternative ist haltbar. Die erste untergräbt sich selbst; denn sobald man bemerkt, daß die eigene Überzeugung ihre Wahrheit bedingt, kann man sie nicht länger unqualifiziert aufrecht erhalten.

Änderungen der Überzeugung, daß die Verwirklichung eines Wunsches befriedigt: Wenn wir gesagt haben, daß diese Überzeugung den Wunsch notwendig begleitet, so sollte das nicht heißen, daß sie ein Produkt dieses Wunsches

ist. Im Gegenteil, in der Lesart (b) ist sie eine normale Überzeugung mit normalen Erkenntnisbedingungen. Sie kann durch Erfahrung begründet sein oder nicht; sie kann falsch sein; man kann herausfinden, daß sie falsch ist; und, um es so allgemein zu sagen, daß auch der probabilistische Fall abgedeckt ist, sie kann sich auf der Grundlage neuer Erfahrungen entsprechend den allgemeinen Regeln rationaler Überzeugungsänderung festigen oder verlieren. Welche Folgen hat eine solche Änderung für die Wünsche? Auch hier sind wieder zwei Fälle zu unterscheiden.

Der Normalfall ist, daß eine Überzeugungsänderung nur eine Neuberechnung der erwarteten Nutzenwerte nach sich zieht. Das kann sogar dann der Fall sein, wenn sich eine Annahme über zukünftige Befriedigung verändert. Zum Beispiel bestelle ich in einem Restaurant ein exotisches und mir völlig unbekanntes Gericht. Ich probiere es, es schmeckt mir nicht besonders, und so revidiere ich meine Überzeugung über die Befriedigungswirkung dieses Gerichts und demgemäß auch seinen erwarteten Nutzen. Dieser Fall läßt sich aber nur deshalb so einfach beschreiben, weil der Zirkel hier nicht auftritt; in der Regel beeinflußt eben der Wunsch nach einem Gericht nicht, wie es schmecken wird.

Ganz anders verhält es sich, wenn der Überlegungszirkel auftritt. Die Änderung einer Überzeugung über zukünftige Befriedigung, d. h. der in (3) involvierten Wahrscheinlichkeiten, erzeugt dann eine schwierige Lage; der kontrollierende Wunsch nach Befriedigung und die kontrollierten Wünsche nach anderen Dingen sind dann aus dem Gleichgewicht und lassen sich rationalerweise nicht länger aufrechterhalten. Irgendwie sind sie zu ändern, irgendwo muß ein neues Gleichgewicht gefunden werden; doch ist ganz unklar, wie man sie ändern und wo das Gleichgewicht zu finden sein soll. Jeder kennt die Erfahrung des Leerlaufs – wenn z. B. die Faszination des Rock'n'roll verblaßt oder eine Liebe verschlissen ist – und den oft quälenden Prozeß einer Neuausrichtung der eigenen Wünsche.

Die Gleichung (3) als Modell der praktischen Überlegung: Die Gleichungen (1) und (2) liefern Regeln des praktischen Überlegens; sie sagen uns, wie man extrinsische und Gesamtnutzenwerte ableitet. Man könnte daher dazu neigen, die Gleichung (3), die im Zirkelfall die Gleichungen (1) und (2) ersetzt, ebenso aufzufassen, nämlich als die Regel: Löse (3); gibt es keine Lösung, so ist nichts zu machen (obgleich dies durch geeignete Stetigkeitsannahmen ausgeschlossen werden könnte); gibt es eine Lösung, so suche die maximale Lösung für $U'(A)$, denn bei ihr ist die erwartete Befriedigung am höchsten! Wäre dies ein durchführbares Überlegungsmodell, so wäre es sinnvoll, die zugehörige mathematische Theorie aus-

zuarbeiten. Aus zwei Gründen erscheint uns dies jedoch äußerst fragwürdig.

Erstens stehen die relevanten Wahrscheinlichkeiten kaum zur Verfügung; sie zu kennen, würde bedeuten, für jeden beliebigen Sachverhalt A auf die folgende Frage eine Antwort zu haben: „Wenn ich A so und so stark wünsche, wie wahrscheinlich ist es dann, daß A mich in dem und dem Maße befriedigen würde?“ Nun kennen wir aber einfach viele Arten von Befriedigung nicht und können sie nicht einstufen. Wie fühlt man sich als Frau von N. N. oder als chinesischer Reisbauer oder als Himalaya-Kletterer? Schwer zu sagen. Vollends hoffnungslos wird es, wenn auch noch der Einfluß der Wünsche zu berücksichtigen ist. Wie befriedigend wäre das Klettern im Himalaya, wenn man es gerne täte? Uns scheint, wir können solche Fragen nur beantworten, wenn es um Arten von Befriedigung geht, die wir schon kennen, und um Typen von Wünschen, die wir schon verfolgt haben; alle anderen Wahrscheinlichkeiten wären pure Raterei.

Zweitens ist die Ausbildung von Zielen und Zwecken, d. h. von nicht abgeleiteten Wünschen, ausgesprochen dunkel. Die Vorstellung, man könnte sich, sozusagen durch schiere Willenskraft, die von einer solchen Überlegung empfohlenen Wünsche zu eigen machen, scheint doch reichlich abwegig. Von daher ist es ausgesprochen zweifelhaft, daß eine solche Überlegung den Einfluß auf die Einstellungen und Handlungen des Subjekts haben kann, aus dem sie erst ihren Sinn bezöge.

Rationale und wahre Wünsche: Wenn die Formel (3) nicht als ein Modell der praktischen Überlegung zu verstehen ist, was leistet sie dann? Sie beschreibt ein Lernmodell, zumindest in dem schwachen Sinn, daß nach ihr Wünsche, die nicht im Gleichgewicht sind, rationalerweise zu verändern sind und daß diese Veränderung sich auf ein neues Gleichgewicht zuzubewegen hat.

Tatsächlich verfügen wir mit (3) über einen reicheren Begriff rationaler Wünsche. Im Rahmen des bisherigen Bildes gab es keine Möglichkeit, intrinsische Wünsche oder Nutzenwerte als rational oder irrational zu qualifizieren; nur für extrinsische Wünsche gab es dafür einen präzisen Sinn. Unser verbessertes Bild enthält jedoch zudem für kontrollierte wie für kontrollierende Wünsche einen Rationalitätsstandard (und eröffnet dadurch die Möglichkeit, einen Begriff von rationalen Gefühlen zu bilden).³² Anders gesagt: (3) erweitert unsere Möglichkeiten zur Kritik von Wünschen, und zwar zu einer Kritik, die sich nicht auf moralische oder

³² Wir denken dabei gerade an den in Anmerkung 30 dargelegten Sachverhalt.

andere externe Standards stützt, sondern sich ausschließlich an die in der Person selbst liegenden Maßstäbe hält.³³

Insofern die Überzeugung, daß die Erfüllung der eigenen Wünsche befriedigen wird, wahr sein kann und die Wahrscheinlichkeiten in (3) objektiv sein können, verleiht (3) auch der Idee wahrer Wünsche genauere Bedeutung: Unsere wahren Wünsche sind jene, die ausgehend von wahren Überzeugungen in einem maximalen Gleichgewicht sind. Wir sind rationalerweise auf der Suche nach unseren wahren Wünschen; da aber Befriedigung auf immer wieder andere Weise zustande kommt, muß die Suche immer wieder von neuem anheben. Die Gleichung (3) ist uns dabei nicht wirklich eine praktische Hilfe, aber sie vermittelt uns eine gewisse theoretische Einsicht in das, wonach wir suchen.

Literatur

- Bentham, J.: *An Introduction to the Principles of Morals and Legislation*. Hg. von J. H. Burns and H. L. A. Hart. London 1970.
- Burge, T.: Individualism and the Mental. *Midwest Studies in Philosophy* IV (1979), S. 73–121.
- Butler, J.: Fifteen Sermons Preached at the Rolls Chapel. In *Butler's Sermons and Dissertation on Virtue*. Hg. von W. R. Matthews. London 1949.
- Campbell, R./Sowden, L. (Hg.): *Paradoxes of Rationality and Cooperation*. Vancouver 1985.
- Churchland, P. M.: The Logical Character of Action-Explanations. *Philosophical Review* 79 (1970) S. 214–236, dt. Übers. in: *Analytische Handlungstheorie*. Band 2: *Handlungserklärungen*. Hg. von A. Beckermann. Frankfurt a. M. 1977, S. 304–331.
- Davidson, D.: *Essays on Actions and Events*. Oxford 1980, dt. Übers.: *Handlung und Ereignis*. Frankfurt a. M. 1985.
- Eells, E.: *Rational Decision and Causality*. Cambridge 1982.
- Elster, J.: *Sour Grapes*. Cambridge 1983.
- Gosling, J. C. B.: *Pleasure and Desire: The Case for Hedonism Reviewed*. Oxford 1969.
- Griffin, J.: *Wellbeing*. Oxford 1986.
- Jeffrey, R. C.: *The Logic of Decision*. Chicago 1965. 2. erw. Ausg. 1983, dt. Übers.: *Logik der Entscheidungen*. München 1967.
- Kant, I.: *Kritik der praktischen Vernunft*. 1788.
- Kusser, A.: *Dimensionen der Kritik von Wünschen*. Frankfurt a. M. 1989.
- Lewis, D.: Desire as Belief. *Mind* 97 (1988), S. 323–332.
- Michael, R. T./Becker, G. S.: On the New Theory of Consumer Behavior. *Swedish Journal of Economics* 75 (1973), S. 373–396.
- Mill, J. St.: *Der Utilitarismus*. Stuttgart 1976.

³³ Dieser Punkt ist in Kusser (1989), S. 185–193, ausführlicher dargelegt.

- Von Neumann, J./Morgenstern, O.: *Theory of the Games and Economic Behavior*. Princeton 1944.
- Pears, D.: *Motivated Irrationality*. Oxford 1986.
- Pigou, A. C.: *The Economics of Welfare*. New York 1920.
- Raiffa, H.: *Decision Analysis*. Reading 1969, dt. Übers.: *Einführung in die Entscheidungstheorie*. München 1973.
- Rawls, J.: *A Theory of Justice*. Cambridge, Mass. 1971.
- Russell, B.: *The Analysis of Matter*. London 1921.
- Samuelson, P. A.: *Foundations of Economic Analysis*. Cambridge 1947.
- Savage, L. J.: *The Foundations of Statistics*. New York 1954, 2. Aufl. 1972.
- Sen, A. K.: Behaviour and the Concept of Preference. *Econometrica* 40 (1973), S. 241–259.
- Sen, A. K.: Plural Utility. *Proceedings of the Aristotelian Society* 81 (1980/81), S. 193–215.
- Sidgwick, H.: *The Methods of Ethics*. 7. Aufl. London 1962.
- Skyrms, B.: *Pragmatics and Empiricism*. New Haven 1984.
- Spohn, W.: *Grundlagen der Entscheidungstheorie*. Kronberg/Ts. 1978.
- Spohn, W.: Stochastic Independence, Causal Independence, and Shieldability. *Journal of Philosophical Logic* 9 (1980), S. 73–99.
- Spohn, W.: Direct and Indirect Causes. *Topoi* 9 (1990), S. 125–145.
- Tugendhat, E.: *Probleme der Ethik*. Stuttgart 1984.
- Tversky, A./Kahneman, D.: The Framing of Decisions and the Psychology of Choice. In: *Rational Choice*. Hg. von J. Elster. New York 1986.