

# Präferenzen, Wohlergehen und Rationalität – Zu den begrifflichen Grundlagen des libertären Paternalismus und ihren Konsequenzen für seine Legitimierbarkeit

ANDREA KLONCHINSKI, REGENSBURG & JOACHIM WÜNDISCH,  
DÜSSELDORF

*Zusammenfassung:* Der libertäre Paternalismus (LP) genießt in Politik, Wissenschaft und Öffentlichkeit eine große Popularität, die er insbesondere zwei Merkmalen verdankt: Erstens stützt sich der LP auf verhaltensökonomische Ergebnisse, die zeigen, dass individuelle Entscheidungen oft nicht der neoklassischen Rationalitätskonzeption entsprechen, sodass Individuen durch sogenanntes *Nudging* zu besseren, ihren wahren Präferenzen entsprechenden Entscheidungen verholfen werden könne. Zweites ist damit der Anspruch verbunden, das Wohlergehen der Individuen, *wie sie selbst es verstehen*, zu erhöhen. Dieser Beitrag zeigt anhand einer dogmengeschichtlichen Analyse der zentralen, dem LP zugrunde liegenden Begriffe Präferenz, Nutzen, Rationalität und Wohlergehen, dass der LP diesen Anspruch nicht einlösen kann. Eine detaillierte begriffliche Analyse ist deshalb geboten, weil diese Konzepte in der Ökonomik immer wieder für Missverständnisse sorgen. Eine theoriegeschichtlich informierte Untersuchung wiederum ist notwendig, da diese Konfusion insbesondere auf der Tatsache beruht, dass der Gehalt dieser Begriffe sich im Laufe der Zeit gravierend verändert hat und sich die verschiedenen Bedeutungen heute z.T. wechselseitig überlagern.

*Schlagwörter:* Libertärer Paternalismus, Nudge, Verhaltensökonomik, bounded rationality, Präferenzen, Theoriegeschichte, Wohlergehen

## 1. Einleitung

Die Debatte um das sogenannte *Nudging* ist spätestens seit Anfang 2015 auch in Deutschland nicht nur in akademischen Kreisen aktuell. Nach dem Vorbild von Großbritannien und den USA gründete zu der Zeit auch die Bundesregierung eine *Nudge*-Abteilung (vgl. Dams et al. 2015). Da „viele Menschen so handeln, dass es ihren eigenen Interessen widerspricht“, so ein Regierungssprecher (zitiert nach Plickert/Beck 2014), könne die Politik hier zum Wohl der Bürger korrigierend eingreifen und diese in die richtige, ihren eigentlichen Interessen entsprechende Richtung „schubsen“. Diese Idee findet ihre theoretische Fundierung im sogenannten libertären Paternalismus (LP) von Richard H. Thaler und Cass R. Sunstein (T&S). Er bezieht seine Legitimation im Wesentlichen aus der verhaltensökonomischen These, dass Individuen häufig nicht entsprechend der neoklassischen Rationalitätskonzeption handeln, sodass ihre offenbarten Präferenzen nicht ihre wahre Präferenzordnung widerspiegeln. Im Bereich der normativen Ökonomik folgt daraus, dass offenbarte Präferenzen nicht mehr als Indikator für Wohlergehen dienen können. Folglich seien politische Maßnahmen gerechtfertigt, die dem Individuum durch *Nudging* zu besseren, seinen eigentlichen Präferenzen entsprechenden Entscheidungen verhelfen. Der damit verbundene Paternalismus werde dadurch abgemildert, dass er die verfügbaren Entscheidungsalternativen nicht einschränke, die Entscheidungsfreiheit der Individuen also unangetastet lasse. Es handele sich somit um einen *libertären* Paternalismus.

Ein Großteil der Sekundärliteratur zum LP beschäftigt sich mit der normativen Frage seiner Legitimierbarkeit und diskutiert, inwiefern er tatsächlich als freiheitsbewahrend (libertär) einerseits und paternalistisch andererseits gelten kann. Der vorliegende Beitrag leistet keine Analyse der potentiell

freiheitsbewahrenden Aspekte des LP, sondern konzentriert sich auf den paternalistischen Aspekt des LP. Er nimmt dabei eine grundlegende Perspektive ein und offeriert eine theoriegeschichtlich informierte Analyse der zentralen Konzepte Rationalität, Präferenz, Nutzen und Wohlergehen sowie ihres Verhältnisses zueinander. Eine solche begriffliche Untersuchung ist deshalb geboten, weil diese Konzepte in der Ökonomik immer wieder für Missverständnisse sorgen. Eine theoriegeschichtlich informierte Perspektive wiederum ist notwendig, da diese Konfusion insbesondere auf der Tatsache beruht, dass der Gehalt dieser Begriffe sich im Laufe der Zeit gravierend verändert hat und sich die verschiedenen Bedeutungen heute z.T. wechselseitig überlagern.

Die Untersuchung gliedert sich wie folgt. Zunächst wird im zweiten Abschnitt die Grundidee des LP dargestellt, woraufhin sich der dritte Abschnitt einer theoriegeschichtlichen Betrachtung der Rationalitätskonzeption bzw. des Verhältnisses von Präferenzen und Nutzen in der positiven Ökonomik zuwendet. Dabei wird herausgearbeitet, dass die frühen Neoklassiker einen hedonistischen Nutzenbegriff vertraten, die moderne Ökonomik jedoch auf einem formalen Nutzenkonzept basiert, welches dazu dient, Präferenzordnungen zu repräsentieren. Was ein Individuum *warum* präferiert und ob das Gewählte letztlich sein Wohlergehen befördert, spielt im Rahmen der modernen Nutzentheorie keine Rolle. Im vierten Abschnitt wird aufgezeigt, dass die Verhaltensökonomik diese Entwicklung zum Teil rückgängig macht, indem sie sich den psychologischen Faktoren hinter beobachtbaren Entscheidungen zuwendet. Die Rationalitätskonzeption der modernen Ökonomik wird dabei jedoch nicht grundsätzlich hinterfragt, sondern bleibt als normatives Ideal individuellen Entscheidens unangetastet. Das verhaltensökonomische Ergebnis, dass Individuen diesem

Ideal oft nicht entsprechen, stellt ein fundamentales Problem für die welfaristische Wohlfahrtsökonomik dar, die Wohlergehen als die Erfüllung konsistenter Präferenzen definiert. Dies ist das sogenannte Vereinbarkeitsproblem und der LP kann als Reaktion auf dieses Problem verstanden werden, wie der fünfte Abschnitt herausstellt. Dabei stellt sich die Frage, welchen Begriff von Wohlergehen der LP an die Stelle der Konzeption von Wohlergehen als Präferenzenerfüllung setzt. Diese Frage wird vor dem Hintergrund der dogmengeschichtlich-systematischen Überlegungen diskutiert. Dass die Wohlergehenskonzepation des LP nicht überzeugt und dass T&S den Individuen letztlich eine bestimmte Vorstellung davon, was gut für die Menschen sei, aufzotroyieren, ist dabei auf in der wirtschaftswissenschaftlichen Dogmengeschichte liegende Gründe zurückzuführen. Das Fazit fasst die Ergebnisse der Untersuchung zusammen und stellt heraus, dass sich der LP nicht auf die Interessen der Betroffenen stützen kann. Somit scheitert der LP an seinen eigenen Ansprüchen, unabhängig davon, ob er tatsächlich libertär ist bzw. ob Paternalismus generell libertär sein kann.

## 2. Der libertäre Paternalismus: die Grundidee

Ausgangspunkt von T&S (2009: 5f.) ist die Einsicht, dass Menschen häufig schlechte Entscheidungen treffen. So ziehen sie Süßigkeiten Früchten vor, obwohl sie eine Diät machen, rauchen, obwohl sie eigentlich aufhören wollten, und legen zu wenig für das Alter zurück, obwohl sie dies als wichtig erachten. Derartige alltägliche (Selbst-)Beobachtungen sowie Ergebnisse der Verhaltensökonomik lassen darauf schließen, dass konkrete Entscheidungen kaum oder gar nicht auf Basis einer vorab bestehenden, konsistenten und über die Zeit stabilen Präferenzordnung erfolgen, sondern sehr stark von der jeweiligen

Entscheidungssituation beeinflusst sind: „What they [people] choose is strongly influenced by details of the context in which they make their choice, for example default rules, framing effects [...], and starting points“ (Sunstein/Thaler 2003: 1161). Daher gewinnt die sogenannte Entscheidungsarchitektur, also die Darstellung einer Entscheidungssituation, an Relevanz (vgl. T&S 2009: 3). Unter der Prämisse, dass die Wahl irgendeiner Entscheidungsarchitektur ohnehin unvermeidbar sei, fordern T&S (2009: 6), der jeweilige Entscheidungsarchitekt solle die Entscheidungssituation derart präsentieren, dass die Individuen bessere Entscheidungen trafen; solche also, die sie gefällt hätten „if they had paid full attention and possessed complete information, unlimited cognitive abilities, and complete self-control“.<sup>1</sup>

In der These, dass es legitim für den privaten oder öffentlichen Entscheidungsarchitekten sei, das Verhalten der Individuen dergestalt zu beeinflussen, dass es ihr Leben „longer, healthier, and better“ mache, liegt das paternalistische Element des LP, welches T&S (2009: 5) selbst wie folgt definieren: „In our understanding a policy is paternalistic if it tries to influence choices in a way that will make choosers better off, as judged by themselves.“<sup>2</sup> Sie betonen, dass es sich bei dem von ihnen propagierten Paternalismus um eine Form des schwachen („*weak*“) Paternalismus handele, da er nicht mit Verboten ar-

---

1 Als paradigmatisches Beispiel für eine Entscheidungsarchitektur führen T&S (2009: 1f.) das Beispiel des Produktarrangements in einer Schulcafeteria an und weisen darauf hin, dass sich durch die Anordnung der Produkte in erheblichem Maße beeinflussen lasse, ob die Kinder eher zu Süßigkeiten oder zu Obst greifen. Dem LP gemäß gelte es dann, die verfügbaren Lebensmittel mit diesem Ziel zu positionieren: „to make the students best off, all things considered“ (T&S 2009: 2). Darauf wird in Abschnitt 5 zurückzukommen sein.

2 Vgl. Sunstein/Thaler (2003: 1162) und T&S (2003: 175).

beite, die verfügbaren Entscheidungsalternativen mithin nicht einschränke und die Entscheidungsfreiheit der Individuen somit unangetastet lasse (vgl. Sunstein/Thaler 2003: 1162; T&S 2009: 6).<sup>3</sup> Ihre libertäre Grundhaltung drücken T&S (ebd.) folgendermaßen aus: „in general, people should be free to do what they like – and to opt out of undesirable arrangements if they want to do so. [...] We strive to design policies that maintain or increase freedom of choice.“ An die Stelle von Verboten treten im Rahmen des LP *Nudges*, d.h. „any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid“ (ebd.). Zudem sind *Nudges* durch eine spezifische Wirkungsweise gekennzeichnet. So charakterisieren T&S (2009: 9) sie als „any factor that significantly alters the behavior of Humans, even though it would be ignored by Econs. Econs respond primarily to incentives. [e.g. taxing candy] [...] Humans respond to incentives, but they are also influenced by nudges.“<sup>4</sup> Dieser Punkt bedarf der näheren Erläuterung.

Als „Econs“ bezeichnen T&S (2009: 7) die imaginären Protagonisten ökonomischer Lehrbücher, die über vollständige Information, Aufmerksamkeit, Selbstkontrolle und unbegrenzte kognitive Fähigkeiten verfügen und somit grundsätzlich im-

---

3 Für eine kritische Auseinandersetzung mit dieser Paternalismusdefinition vgl. Hausman/Welch (2010: 126ff.). Es sei darauf hingewiesen, dass nicht alle von T&S vorgeschlagenen Maßnahmen überhaupt als paternalistisch bezeichnet werden können. So diskutieren sie etwa Maßnahmen zur Erhöhung des Organaufkommens, zum Stoppen des Klimawandels oder zu einer Verringerung der Umweltverschmutzung.

4 Es sei angemerkt, dass die zahlreichen und sehr disparaten Beispiele für *Nudges*, die T&S liefern, nicht alle den genannten Definitionen entsprechen, worauf auch Hausman/Welch (2010: 125) und Selinger/Whyte (2012: 11) hinweisen.

mer die bestmöglichen Entscheidungen treffen. Sie lassen sich insbesondere nicht von der Gestaltung, dem *framing* der Entscheidungssituation beeinflussen. „Humans“ hingegen sind echte Menschen, die gewöhnlich nicht über das genannte Ausmaß an Kenntnissen und Fähigkeiten verfügen und deren Entscheidungen von verschiedenen *biases* verzerrt werden könnten. Das liegt daran, dass Econs niemals Entscheidungen treffen, ohne ihr „reflexives System“ zu konsultieren, während Humans sich häufig auf ihr „automatisches“ und manchmal fehlerhaftes System verlassen (vgl. T&S 2009: 24). Unter reflexivem Denken wird das konzentrierte, bewusste, oft beschwerliche Nachdenken über ein Problem verstanden. So ist das reflexive System aktiv, wenn wir komplexere mathematische Aufgaben lösen (was ist  $24^2 \times 130$ ?) oder über die Strukturierung eines wissenschaftlichen Artikels nachdenken. Das automatische System hingegen arbeitet schnell, unbewusst und mühelos. Gelernte Tätigkeiten, wie Reiten, Klettern oder Fahrradfahren, werden durch dieses System gesteuert ebenso wie die Lösung einfacher Aufgaben (was ist  $1 + 1$ ?). Während automatisches Denken für den Alltag unerlässlich ist, ist es gerade aufgrund seines automatischen, unkontrollierten und unreflektierten Funktionierens anfällig für Fehler, wie die Verhaltensökonomik ausführlich demonstriert.<sup>5</sup> *Nudges* zielen auf die Korrektur derartiger Fehler des automatischen Systems ab.<sup>6</sup> Der LP führt die Individuen damit hin zu den Entscheidungen, die sie selbst getroffen haben

---

5 Diese Differenzierung entspricht der Unterscheidung von schnellem und langsamem Denken von Kahneman (2011). Siehe auch Ross (2013: 204–11).

6 Insofern *Nudges* damit auf einer unbewussten Ebene wirken, über die die Individuen wenig bis keine Kontrolle haben (vgl. Kahneman 2011: 22), sind sie in der Literatur als manipulativ kritisiert worden. Siehe Grüne-Yanoff (2012: 636), Coons/Weber (2013: 22) und White (2013: 95f.).

würden, wenn sie auf Basis ihres reflexiven Systems und somit konsistenten Präferenzen entsprechend gehandelt hätten. Was aber sind die „wahren“ Präferenzen des Individuums und warum sollte es überhaupt gut sein, einer konsistenten Präferenzordnung gemäß zu handeln? Dies sind grundsätzliche Fragen, von denen letztlich die Legitimierbarkeit des LP abhängt. Zu ihrer Beantwortung gilt es, das Rationalitätskonzept in seiner historischen Entwicklung näher zu betrachten.

### 3. Präferenzen in der positiven Ökonomik

Die Bedeutung der Begriffe Präferenz, Nutzen und Rationalität zueinander unterlagen in der Geschichte der Wirtschaftswissenschaften einem erheblichen Wandel.<sup>7</sup> Um die verschiedenen Bedeutungsdimensionen, die heute oft nur unzureichend voneinander differenziert werden (vgl. Broome 1991a: 65f., 1991b: 4), herauszuarbeiten, muss die Untersuchung am Ende des neunzehnten Jahrhunderts ansetzen. Im Zuge der sogenannten „marginalen Revolution“ der Ökonomik von 1871 rückte das Verhalten des individuellen Akteurs erstmals in das Zentrum des wirtschaftswissenschaftlichen Interesses (vgl. Ross 2014: 29–33). Als marginale Revolution bezeichnet man die Einführung der Differentialrechnung in die Ökonomik und damit die „Entdeckung“ des Prinzips des abnehmenden Grenznutzens durch William S. Jevons, Carl Menger und Léon Walras. Im Zentrum dieser Revolution steht die subjektive Werttheorie. Während die klassischen Ökonomen zur Erklärung der Entstehung von Preisen die Seite der Güterproduktion betrachteten,

---

7 Die folgende Darstellung beruht auf Klonschinski (2016: Kap. 2) und findet sich ähnlich auch bei Klonschinski (2013).

fokussierten die Marginalisten die Nachfrageseite und warfen damit die Frage auf, wie Wirtschaftssubjekte Entscheidungen treffen (vgl. Blaug 1997: 277ff.). Die Erklärung der Entstehung und Natur von Nachfrage bedurfte bestimmter Annahmen über das Verhalten der Konsumenten und zu diesem Zweck wurde auf den hedonistischen Nutzenbegriff rekurriert (vgl. Mandler 1999: 68).

So stellt Jevons (1911/2006: 1f.) in seiner „Theory of Political Economy“, die als exemplarisch für das frühe neoklassische Denken gelten kann (vgl. Ross 2014: 39), fest: „Repeated reflection and inquiry have led me to the somewhat novel opinion, that [sic] value depends entirely upon utility. [...]. [W]e have only to trace out carefully the natural laws of the variation of utility, as depending upon the quantity of commodity in our possession, in order to arrive at a satisfactory theory of exchange.“ „Utility“ versteht er dabei in Anlehnung an Jeremy Bentham in einem hedonistischen Sinne: „Utility must be considered as measured by, or even as actually identical with, the addition made to a person’s happiness. It is a convenient name for the aggregate of the favourable balance of feeling produced – the sum of the pleasure created and the pain prevented“ (Jevons 1911: 45). Hinsichtlich der Frage, wie genau die „natural laws of utility“ aussehen, ist besagte Entdeckung des Prinzips vom abnehmenden Grenznutzen von entscheidender Bedeutung. So stellt Jevons fest, dass nicht jede zusätzliche Einheit eines Gutes denselben Nutzen, d.h. dieselbe Befriedigung stifte (vgl. ebd.: 44). Vielmehr werde der Grenznutzen, also der Nutzen, den eine zusätzliche Einheit eines Gutes stifte, umso kleiner, je größer die Gesamtmenge des bereits konsumierten Gutes sei (vgl. ebd.: 53). Zur Begründung dieses Prinzips rekurriert Jevons auf die Psychophysik, die das Verhältnis physischer Stimuli und subjektiver Empfindung unter Verwendung experimentel-

ler Methoden untersucht.<sup>8</sup> Die Empfindung, so die Annahme, könne indirekt durch die Menge des Stimulus gemessen werde, die nötig sei, um eine gerade noch spürbare Veränderung in der Intensität des Sinneseindrucks zu generieren. Es wird deutlich, dass am Ende des neunzehnten Jahrhunderts eine enge Verbindung zwischen der sich entwickelnden psychologischen und der ökonomischen Disziplin bestand.<sup>9</sup> Im Hinblick auf die nachfolgende Entwicklung dieses Verhältnisses sowie der Verhaltensökonomik verdient diese Tatsache besondere Beachtung.

Nachdem Jevons (1911: 59) den Grenznutzen als entscheidende Variable zur Erklärung des Verhaltens von Wirtschaftsakteuren identifiziert hat, postuliert er, es läge der menschlichen Natur eine unvermeidliche Tendenz zugrunde, jeweils gemäß dem größtmöglichen Vorteil zu entscheiden. Wenn eine Person daher mit einer gewissen Verteilung ihrer Ressourcen auf verschiedene Verwendungsweisen zufrieden sei, ließe sich daraus schließen, dass keine Umverteilung ihr zusätzlichen Nutzen stiften würde. Der Grenznutzen der beiden Verwendungen müsse demzufolge gleich sein. „The general result“, so folgert Jevons (1911: 60), „is that commodity, if consumed by a perfectly wise

---

8 Zur Relevanz der Psychophysik für Jevons siehe Maas (2005: 164ff.), Bruni/Sugden (2007: 150f.), Ross (2013: 39) und Heukelom (2014: 2). Stärker noch als Jevons stützte Ysidro Edgeworth seine ökonomischen Überlegungen auf die Psychophysik. Er betrachtete „man as a pleasure machine“ und ersann ein Hedonimeter als Messgerät: „To precise the ideas let there be granted to the science of pleasure what is granted to the science of energy; to imagine an ideally perfect instrument, a psychophysical machine, continually registering the height of pleasure experienced by an individual“ (Edgeworth 1881: 15, 101). Vgl. Schabas (1990: 123f.) und Bruni/Sugden (2007: 151–3).

9 Schabas (2007: 137) verdeutlicht dieses Verhältnis wie folgt: „economics was firmly rooted in psychology, which was something to relish, not fear. It would bring greater rigor and objectivity to the subject and provide the proper ontological foundation for the newly developed mathematical theory.“

being, must be consumed with a maximum production of utility“. Analoges gelte für den Tausch am Markt. Dieser würde so lange fortgesetzt, bis keine der beiden Parteien durch weitere Tauschhandlungen ihre Freude vergrößern kann (vgl. ebd.: 86, 96ff.); m.a.W., bis ihr individueller Nutzen jeweils maximal ist.

Zusammenfassend lässt sich festhalten, dass Jevons den utilitaristischen Nutzenbegriff als ein Instrument zur Analyse des Nachfrageverhaltens der Konsumenten in die Ökonomik einführte, mit dessen Hilfe er ökonomische Phänomene wie die Entstehung von Marktpreisen auf *psychologische* Faktoren zurückführen konnte (vgl. Schabas 2007: 14, 134ff.). Es gilt in dem hier interessierenden Kontext zu beachten, dass das Erzielen von Nutzen im Rahmen der hedonistischen Nutzentheorie Ziel und Motiv der individuellen Handlungen ist. Die Individuen agieren in einer bestimmten Art und Weise, *weil* sie ihren Nutzen in Form von angenehmen mentalen Zuständen maximieren wollen.

Die frühe neoklassische Theorie, insbesondere ihre Annahme des psychologischen Hedonismus, geriet Anfang des 20. Jahrhunderts nicht zuletzt aufgrund der Verbreitung positivistischer Ideen zunehmend in die Kritik. Lionel Robbins „An Essay on the Nature & Significance of Economic Science“ gilt dabei als Meilenstein. Robbins strebte nach einer klaren Abgrenzung zwischen der positiven Ökonomik als *reiner* Wissenschaft auf der einen Seite und normativen Aussagen, die aufgrund ihrer Werthaltigkeit keine wissenschaftliche Basis hätten, andererseits. Wolle die Ökonomik eine Wissenschaft sein, so habe sie sich auf die Untersuchung beobachtbarer Phänomene zu beschränken. Da die spezifischen Gründe, die hinter den Entscheidungen von Wirtschaftssubjekten stehen, nicht beobachtbar seien, können über sie auch keine wissenschaftlichen Aussagen gemacht werden. Dies, so Robbins (1935: 86), sei jedoch unproblematisch,

denn diese Gründe seien für die Ökonomik ohnehin irrelevant: „*Why the human animal attaches particular values [...] to particular things, is a question which we do not discuss.*“ Um eine Werttheorie formulieren zu können, müsse lediglich angenommen werden, „*that individuals can arrange their preferences in an order, and in fact do so*“ (ebd.: 78f.). Präferenzordnungen haben dabei den Vorteil, dass sie beobachtet werden können und somit wissenschaftlicher Analyse zugänglich sind.

Bereits Vilfredo Pareto (1927/1971: 113) hatte gezeigt, dass das Wissen um die Menge des Nutzens i.S. der Befriedigung, die ein Individuum aus dem Konsum eines Gutes zieht, für die Werttheorie irrelevant sei. Entscheidend sei lediglich die Kenntnis der jeweiligen Kombination zweier Güter, zwischen denen das Individuum indifferent sei. Geometrisch lassen sich diese Punkte mithilfe sogenannter Indifferenzkurven abbilden (vgl. ebd.: 119). Unter der Annahme, dass das Individuum eine insgesamt höhere Menge an Gütern einer geringeren vorzieht, zeigen im Diagramm nach rechts oben verschobene Kurven einen höheren Nutzen an und lassen sich mit einem höheren Nutzenindex versehen. Welche Zahlen *genau* den Indifferenzkurven zugewiesen werden, ist dabei arbiträr, solange sie nur die *Präferenzordnung* des Individuums widerspiegeln. Anstelle eines kardinalen Nutzenmaßes, welches die Intensität der empfundenen Freude widerspiegelt, tritt somit ein ordinales Nutzenmaß, welches lediglich die Rangordnung der Alternativen angibt. Pareto kann somit als Vorläufer der von Robbins eingeläuteten „ordinalen Revolution“ gelten. Diese Ideen Paretos nimmt John Hicks in seiner Formulierung der modernen Nachfragetheorie auf. Damit die Präferenzen der Konsumenten sich jedoch in Form einer ordinalen Nutzenfunktion darstellen lassen, müssen sie die formale Bedingung der Konsistenz erfüllen, also vollständig und transitiv sein (vgl. Hicks 1946: 18ff.).

Um ihre Funktion im Rahmen der Nachfragetheorie erfüllen zu können, müssen Präferenzen zudem stabil sein – andernfalls ließe sich jede Änderung des Nachfrageverhaltens als eine Präferenzänderung interpretieren. Diese Überlegungen läuten die Axiomatisierung der Nutzentheorie ein (vgl. Broome 1991b: 3).

Parallel fand eine andere für die Ökonomik zentrale theoretische Entwicklung statt, nämlich Paul Samuelsons Formulierung der *revealed preference theory* (RPT). Samuelson (1947: 3f.) zielte darauf ab, die ökonomische Analyse ausschließlich auf operationalisierbare Theoreme zu beschränken, die sich durch empirische Beobachtungen widerlegen lassen. Die ordinale Nutzentheorie war seiner Meinung nach in dieser Hinsicht nicht weit genug gegangen, da sich nicht nur ein hedonistisch verstandener Nutzenbegriff, sondern auch die Konzepte der Indifferenz und der Präferenz auf innere, nichtbeobachtbare Phänomene beziehen (vgl. Samuelson 1938: 61f.). Samuelson versuchte zu zeigen, dass die Konsistenz beobachtbarer Entscheidungen für die Nachfragetheorie ausreicht. Sind diese Entscheidungen konsistent, so kann dem Individuum nachträglich eine Präferenzordnung zugeordnet werden, die die Entscheidungen repräsentiert (vgl. Hausman 2012: 26). Es ist dabei zu betonen, dass Samuelson in der ursprünglichen Formulierung der Theorie gar nicht von RPT spricht, da er nicht davon ausgeht, dass es irgendetwas zu „offenbaren“ gibt (vgl. Giocoli 2003: 104). Präferenzen werden vielmehr als Entscheidungen *definiert*. In einem späteren Paper gibt Samuelson (1948) diese totale Beschränkung auf beobachtbare Phänomene jedoch auf und geht nunmehr davon aus, dass sich Präferenzen in Entscheidungen *offenbaren* – eine Annahme, die auch Hicks macht (vgl. Davis 2003: 30, Giocoli 2003: 103ff.).<sup>10</sup> Obwohl sich

---

<sup>10</sup> Samuelson (1948: 149) schreibt: „By comparing costs of different combinations of goods at different relative price situations, we can infer

Samuelsons RPT letztlich als äquivalent mit der ordinalen Nutzentheorie erweist, ist ihr Einfluss auf die Ökonomik kaum zu überschätzen (vgl. Giocoli 2003: 106). Insbesondere verstärkte sie zum einen den „informellen Behaviorismus“ (Mandler 1999: 88) und somit das wissenschaftliche Image der Wirtschaftswissenschaft; zum anderen führte die RPT zu einem stärkeren Fokus auf die Repräsentation von Entscheidungen und damit einhergehend zu einer weiteren Abwendung vom Ziel, individuelles Handeln zu *erklären* (vgl. Giocoli 2003: 103, 110).

Während die – behavioristisch inspirierte – ordinale Nutzentheorie nach wie vor die Basis der modernen Nachfrage- und Wohlfahrtstheorie sowie den Kern der ökonomischen Rationalitätskonzeption darstellt (vgl. Hausman 2012: 13), erfordert die Analyse von Entscheidungssituationen unter Risiko und Unsicherheit einen anspruchsvolleren, kardinalen Nutzenbegriff. Hier ist insbesondere der Beitrag von von Neumann und Morgenstern zu nennen. Sie erweiterten die ordinale Nutzentheorie zur sogenannten Erwartungsnutzentheorie (ENT). Damit die Konstruktion einer Nutzenfunktion auch bei Entscheidungen unter Risiko möglich ist, muss die Funktion Präferenzen in Bezug auf verschiedene Wahrscheinlichkeitsverteilungen über die Einzelereignisse abbilden können. Die Präferenzen müssen dann neben der Forderung nach Vollständigkeit und Transitivität einige zusätzliche, stärkere Axiome, wie etwa das Unabhängigkeitsaxiom erfüllen (vgl. von Neumann/Morgenstern 1953: 26ff.).<sup>11</sup> Ist dies der Fall, lässt sich die entsprechende Präferenzordnung durch einen kardinalen Nutzenbegriff repräsentieren, sodass

---

whether a given batch of goods is preferred to another batch; the individual guinea-pig, by market behaviour, reveals his preference pattern – if there is such a consistent pattern.“

11 Siehe zu den Axiomen der ENT bspw. Broome (1991a: 90ff.) und Hausman (2012: 37–42).

ein rationales Individuum per definitionem seinen Erwartungsnutzen maximiert.<sup>12</sup> Es sei angemerkt, dass die Axiome im Allgemeinen als normative Anforderung an rationale Präferenzen verstanden werden (vgl. Broome 1991a: 91).<sup>13</sup> Die ENT stellt also eine *normative* Konzeption individuell rationalen Verhaltens dar und kann somit auch als Rationalitäts*ideal* bezeichnet werden, welches bestimmte Anforderungen an rationale Individuen stellt und rasch zum Standardmodell avancierte (vgl. Heukelom 2014: 96f.).

Obwohl sich die Ansätze im Detail unterscheiden, gilt sowohl für die ordinale Nutzentheorie als auch die ENT, dass der Nutzenbegriff jeweils lediglich dazu dient, konsistente Präferenzordnungen zu repräsentieren. Ein Individuum wählt eine bestimmte Option damit nicht, *weil* diese ihm einen höheren Nutzen verspricht, sondern es ist umgekehrt so, dass die Option im Nachhinein einen Nutzenwert zugeordnet bekommt, weil sie gewählt wird. Individuelle Nutzenmaximierung bedeutet dann nichts weiter als diejenige Option zu wählen, die in der konsistenten und über verschiedene Entscheidungskontexte hinweg stabilen Präferenzordnung oben steht (vgl. Hausman/McPherson 2008: 235f.). „Nutzen“ ist somit zu einem rein formalen Konzept geworden und hat mit dem Motiv der Handlung bzw. daraus resultierenden angenehmen Gefühlen nichts mehr zu tun (vgl. Blaug 1997: 337ff.). Mit der Axiomatisierung der Nutzentheorie wird das Ziel der frühen neoklassischen Nachfrage Theorie, zu erklären, *warum* ein Akteur ein bestimmtes

---

12 Dieser Nutzenbegriff darf nicht mit dem der frühen Marginalisten verwechselt werden. Siehe für eine genaue Differenzierung Ellsberg (1954: 530ff.).

13 Tatsächlich ist die Frage, ob die Präferenzaxiome normativ oder deskriptiv zu verstehen sind, komplexer als hier dargestellt. Siehe dazu ausführlich Heukelom (2014).

Gut wählt, zugunsten der Beschreibung von Verhalten fallengelassen (vgl. Davis 2003: 31). Die skizzierte Entwicklung der Ökonomik lässt sich somit etwas vereinfachend als eine kontinuierliche „Entleerung“ ökonomischer Theorien und Begriffe von psychologischen Annahmen betrachten – Giocoli (2003: 30) spricht von einer „Flucht aus der Psychologie“.<sup>14</sup> Die Verhaltensökonomik nimmt diese Entwicklung z.T. zurück, wie der folgende Abschnitt zeigt.

#### 4. Verhaltensökonomik, Wohlfahrtsökonomik und das Vereinbarkeitsproblem

Die Verhaltensökonomik formierte sich in den 1970er und 1980er Jahren vor dem Hintergrund der sogenannten kognitiven Revolution der Psychologie und dem damit einhergehenden Aufstieg der Kognitionswissenschaft.<sup>15</sup> Lag der Fokus zuvor auf beobachtbarem Verhalten, so rückten nun die dieses Verhalten verursachenden mentalen Prozesse ins Zentrum des Interesses; es ging fortan darum, die „Black Box“ des Geistes zu öffnen (Heukelom 2014: 97). Dementsprechend zielt die Verhaltensökonomik darauf ab, systematische Abweichungen des tatsächlichen individuellen Entscheidens vom neoklassischen Rationalitätskonzept aufzudecken (vgl. Sugden 2008: 227). Mittlerweile weithin bekannte Ergebnisse verhaltensökonomischer Experimente umfassen z.B. *loss aversion*, *status quo bias* und *framing*-Effekte. Vor dem Hintergrund der vorherigen Ausführungen ist zu betonen, dass hier auf die hinter den

---

14 Eigene Übersetzung. Vgl. auch Davis (2003: 30), Bruni/Sugden (2007) und Ross (2013: 196).

15 Vgl. Güth/Kliemt (2013: 3), Ross (2013: 199) und Heukelom (2014: 97), welcher eine ausführliche und detaillierte Analyse der Geschichte der Verhaltensökonomik offeriert.

jeweiligen Entscheidungen stehenden Gründe rekuriert wird.<sup>16</sup> Entsprechend betonen Camerer/Loewenstein (2004: 3): „At the core of behavioral economics is the conviction that increasing the realism of the psychological underpinnings of economic analysis will improve the field of economics.“ Wie das Zitat bereits andeutet, speist sich dieses Forschungsinteresse keineswegs aus einer fundamentalen Kritik an der ENT bzw. der Mainstream-Ökonomik. Es geht vielmehr darum, diese auf Basis experimentell gewonnener psychologischer Daten zu verbessern. Die grundlegende Annahme des nutzenmaximierenden, rationalen Verhaltens bleibt dabei erhalten, die ENT fungiert mithin weiterhin als normatives Ideal (vgl. Güth/Kliemt 2013; Heukelom 2014: 170ff.).

Anfangs im Wesentlichen auf die Finanzökonomik beschränkt, entwickelte die Verhaltensökonomik in den 1990er und 2000er Jahren ein begriffliches Instrumentarium, welches ihr einen breiteren Anwendungsbereich erschloss (vgl. Heukelom 2014: 171–179). Eine wichtige Rolle spielte dabei der teils aus der Psychophysik stammende Ansatz der zwei Systeme, wie er auch dem LP zugrunde liegt. Heukelom (ebd.) zufolge wurde dieser Ansatz im Rahmen der verhaltensökonomischen Erforschung dynamisch inkonsistenter Präferenzen populär. Diese Präferenzen ließen sich als konfligierende Präferenzen von zwei Teilen des Selbst interpretieren: „Thus, the behavioral economists’ way of dealing with intertemporal choice began to describe human behavior as the outcome of two systems or processes striving for dominance“ (ebd.: 176).<sup>17</sup> Die Dichotomie von normativem, nur dem Versuchsleiter bekannten, Rationalitätsideal der ENT auf der einen und dem deskriptiven, von diesem

---

16 Siehe dazu Güth/Kliemt (2013).

17 Siehe dazu auch Sugden (2008: 235f.).

Muster abweichenden Verhalten der Versuchspersonen auf der anderen Seite wurde somit *in das Individuum hinein* verlegt. Systematische Verletzungen des Ideals wurden fortan als ein Versagen des Individuums, sein automatisches System zu kontrollieren, verstanden (vgl. ebd.: 178). Wenn das Individuum in dieser Hinsicht scheitert, stellt sich die Frage, ob der Staat Abhilfe schaffen kann. Die skizzierte Entwicklung legte somit den Grundstein für den LP.

Tatsächlich kam die Frage, wie der normative Zweig der Wirtschaftswissenschaften, die Wohlfahrtsökonomik, mit verhaltensökonomischen Ergebnissen umzugehen habe, erst zu einem Zeitpunkt auf, als die Verhaltensökonomik bereits Einzug in die positive Mainstream-Ökonomik gefunden hatte.<sup>18</sup> Um zu verstehen, warum systematische Abweichungen vom Rationalitätsideal für die Wohlfahrtsökonomik ein Problem darstellen, ist nochmals eine historisch-systematische Betrachtung der Begriffe Präferenz und Wohlergehen angezeigt. Anders als im vorherigen Abschnitt geht es hier jedoch um die spezifische Verwendung dieser Begriffe innerhalb der Wohlfahrtsökonomik. Denn obwohl die Diskurse der positiven und der normativen Ökonomik nicht isoliert voneinander verliefen, so sind sie doch sehr unterschiedlich, wie die folgenden Überlegungen, die uns nochmals an den Anfang des zwanzigsten Jahrhunderts zurückführen, illustrieren.

Als Begründung der Wohlfahrtsökonomik gilt Arthur C. Pigous „The Economics of Welfare“. Pigou (1932/1962: 10) zufolge ist es genuine Aufgabe der Ökonomik, der Politik praktische Anweisungen an die Hand zu geben, wie diese das ökonomische Wohlergehen der Menschen befördern könne. Unter ökonomischem Wohlergehen versteht er dabei den Teil

---

18 Vgl. Sugden (2009: 366), Berg/Gigerenzer (2010: 159), McQuillin/Sugden (2012: 554), und White (2013: 28).

des allgemeinen Wohlergehens, der wissenschaftlichen Methoden zugänglich, also messbar sei: „The one obvious instrument of measurement available in social life is money. Hence, the range of our inquiry becomes restricted to that part of social welfare that can be brought directly or indirectly into relation with the measuring-rod of money. This part of welfare may be called economic welfare“ (ebd.: 12). Das aggregierte Volkseinkommen könne folglich als Indikator der ökonomischen Wohlfahrt einer Nation betrachtet werden (vgl. ebd.: 31). Aufgrund des abnehmenden Grenznutzens von Einkommen ließe sich durch eine Einkommensumverteilung von reich zu arm nun eine Erhöhung der allgemeinen Wohlfahrt bewirken (vgl. ebd.: 89). Den Ärmeren erlaube dieser Einkommenszuwachs die Befriedigung von „more intense wants [...] at the expense of less intense wants“, sodass die aggregierte Summe an Befriedigung (*satisfaction*) steige (ebd.). Wie im vorherigen Abschnitt dargestellt, wurde der hedonistische, kardinale Nutzenbegriff Anfang des zwanzigsten Jahrhunderts scharf kritisiert. So bemängelte Robbins (1935: 91) insbesondere, dass Pigou suggeriere, sein Umverteilungspostulat folge aus der positiven Theorie. Dies sei jedoch nicht der Fall; vielmehr unterstelle Pigou einen identischen Grenznutzen einer zusätzlichen Einkommenseinheit bei verschiedenen Personen und setze somit die Möglichkeit interpersoneller Nutzenvergleiche voraus, was in Robbins Augen ein illegitimes Werturteil darstellt.

Robbins Kritik führte zur sogenannten neuen Wohlfahrtsökonomik, die durch die Ablehnung interpersoneller Nutzenvergleiche charakterisiert ist (vgl. Hausman/McPherson 2008). Um auch ohne derartige Vergleiche Aussagen über die aggregierte Wohlfahrt einer Gesellschaft machen zu können, rekurriert sie auf das sogenannte Pareto-Kriterium, welches besagt, dass eine politische Maßnahme dann vorzugswürdig ist, wenn

sie mindestens eine Person besser, aber niemanden schlechter stellt. Kann eine Verteilung nicht mehr geändert werden, ohne jemanden schlechter zu stellen, handelt es sich um eine Pareto-effiziente Situation. Da Umverteilungen, die niemanden schlechter stellen, in der Realität kaum existieren, formulierten Nicholas Kaldor (1939) und Hicks (1939) das Pareto-Kriterium zum sogenannten Kompensationskriterium um. Demnach ist eine Umverteilung wohlfahrtssteigernd, wenn sie die Profiteure in dem Ausmaß besser stellt, dass diese die Verlierer potentiell kompensieren *könnten* und immer noch besser gestellt wären als vor der Umverteilung.<sup>19</sup>

Inwiefern Individuen „besser“ oder „schlechter“ gestellt werden, wird dabei mit Rekurs auf ihre jeweiligen Präferenzen, operationalisiert in Form von Zahlungsbereitschaften, bestimmt. Die moderne Wohlfahrtsökonomik definiert Wohlergehen damit als die Erfüllung von Präferenzen und ist dezidiert welfaristisch, insofern sie soziale Zustände einzig und allein mit Rekurs auf die individuellen Präferenzen der Betroffenen bewertet (vgl. Sugden 2008: 226; McQuillin/Sugden 2012: 553).<sup>20</sup> In der Akzeptanz individueller Präferenzen als Maßstab für Wohlergehen manifestiert sich dabei die traditionell antipateralistische Ausrichtung der Ökonomik – Ökonomen wollen den Individuen nicht aufoktroieren, was „gut für sie“ ist (vgl. Sugden 2008: 226f.). Dass die Erfüllung von Präferenzen indessen nicht notwendigerweise „gut“ für das jeweilige Individuum ist, liegt auf der Hand – man denke hier z.B. an Drogensucht sowie

---

19 Das Kompensationskriterium ist insbesondere aufgrund der Tatsache kritisiert worden, dass eine Wohlfahrtssteigerung nur die theoretische Möglichkeit, nicht aber die tatsächliche Kompensation der Verlierer fordert. Siehe etwa Reinhardt (1992: 312f.).

20 Zur Konzeption von Wohlergehen als Präferenzenerfüllung siehe Hausman/McPherson (2009).

die Ausbildung von Präferenzen vor dem Hintergrund falscher Überzeugungen oder unzureichender Information – und ist in der Literatur ausführlich diskutiert worden (vgl. Hausman/McPherson 2009; Hausman 2012: 81ff.). Zudem stellt sich auf der Ebene staatlichen Handelns die normative Frage, inwiefern politische Maßnahmen darauf abstellen sollten, individuelle Präferenzen ohne Ansehen ihres Inhalts zu erfüllen. Hier sei etwa auf rassistische, sexistische oder sogenannte teure Präferenzen verwiesen.

Verhaltensökonomische Ergebnisse stellen die welfaristische Wohlfahrtsökonomik jedoch vor ein viel grundlegenderes Problem: Präferenzen können nur dann als Maß für Wohlergehen dienen und die zugeordneten Nutzenwerte oder Zahlungsbereitschaften über Individuengrenzen hinweg aggregiert werden, wenn sie stabil und konsistent sind. Erfüllen Präferenzen diese Anforderungen jedoch nicht, so können sie nicht als Standard für Wohlergehen dienen. Diese Schwierigkeit wird als Vereinbarkeitsproblem von Wohlfahrtsökonomik und Verhaltensökonomik bezeichnet und der LP kann als eine Reaktion auf dieses Problem betrachtet werden (vgl. Sugden 2009: 365; McQuillin/Sugden 2012: 554, 559). Die Frage, wie T&S Wohlergehen definieren, ist Gegenstand des folgenden Abschnitts.

## 5. Libertärer Paternalismus und die Suche nach den „wahren“ Präferenzen

Entscheiden Individuen in bestimmten Kontexten systematisch nicht auf Basis ihres reflexiven Systems, so kann die Entscheidungsarchitektur Abhilfe schaffen und die Betroffenen entsprechend *nudgen* – so die verhaltensökonomisch inspirierte Grundidee des LP. Die zentrale Annahme ist dabei, dass diese *Nudges* zu einer Steigerung des Wohlergehens der Individuen führten

und zwar „as judged by themselves“ (T&S 2009: 5). Dennoch identifizieren die Autoren Wohlergehen keineswegs mit Präferenzenerfüllung (vgl. Sunstein/Thaler 2003: 1163; T&S 2009: 10) – dies würde der Beobachtung, dass Individuen oft fehlerhafte, aus automatischem Denken resultierende Entscheidungen treffen, zuwiderlaufen. Obwohl T&S behaupten, keine gehaltvolle und damit potentiell kontroverse Wohlergehenskonzepktion zu vertreten,<sup>21</sup> basiert ihr Ansatz implizit auf einer Konzeption von Wohlergehen als Erfüllung idealisierter, rationaler Präferenzen.<sup>22</sup> Dies sind diejenigen Präferenzen, die Individuen hätten, wenn sie „unboundedly rational“ wären (McQuillin/Sugden 2012: 560), d.h., „if they had paid full attention and possessed complete information, unlimited cognitive abilities, and complete self-control“ (T&S 2009: 6). Die ENT fungiert somit auch hier als normatives Ideal individuellen Entscheidens, was sich z.B. anhand der Bezeichnung von abweichenden Entscheidungen als „biases and blunders“ manifestiert (ebd.: 19).<sup>23</sup> T&S nehmen damit offenbar an, „that inside every Human there is an Econ – that, deep down, each of us has coherent preferences, of the kind that economic theory has traditionally assumed, and that these can be found by stripping away specific failures of rationality“ (Sugden 2009: 370). Mit dieser Fundierung des Ansatzes in der ENT sowie mit T&Ss spezifischer Wohlergehenskonzepktion sind gravierende Probleme verbunden.

So wurde bislang weder theoretisch noch empirisch gezeigt, dass Personen, die in realen Entscheidungen die Ratio-

---

21 Sunstein/Thaler (2003: 1163) schreiben: „We are not attempting to say anything controversial about welfare, or to take sides in reasonable disputes about how to understand that term.“

22 Vgl. Sugden (2008: 232f.), Qizilbash (2012) und Grüne-Yanoff (2012: 642).

23 Vgl. Sunstein/Thaler (2003: 1168) und T&S (2003: 176).

nalitätsaxiome verletzen, alles in allem ein kürzeres Leben haben, weniger gesund oder weniger glücklich sind (vgl. Berg/Gigerenzer 2010: 148). Die Charakterisierung bestimmter Entscheidungen als „Fehler“ beruht mithin nicht auf empirischen Untersuchungen, die diese Entscheidungen mit negativen Konsequenzen für den Entscheidungsträger in Zusammenhang bringen, sondern lediglich auf den beobachteten Abweichungen vom neoklassischen normativen Ideal (vgl. ebd.). Insofern die ENT eine *normative* Rationalitätskonzeption darstellt, können vermeintliche Fehler im Verhalten jedoch ebenso auf Fehler in der Theorie hinweisen. *Status quo bias* und *loss aversion* etwa könnten sich als fundamentale, durchaus langfristig sinnvolle Charakteristika menschlicher Wünsche herausstellen (vgl. McQuillin/Sugden 2012: 561). Und ob Individuen mit intransitiven Präferenzen außerhalb des Labors jemals wirklich Opfer einer Geldpumpe würden, ist eine offene Frage (vgl. Berg/Gigerenzer 2010: 150).

Ein zentrales Problem der Wohlergehenskonzeption von T&S wurde in der Literatur bereits ausführlich diskutiert, sodass es hier nur kurz genannt werden soll.<sup>24</sup> Es handelt sich dabei um den hohen, stark idealisierten Maßstab, der an rationale Präferenzen angelegt wird. Dabei gälte es zunächst einmal zu definieren, was mit „vollständiger Information“ oder „voller Aufmerksamkeit“ überhaupt genau gemeint ist. Darüber hinaus ist offensichtlich, dass kein menschliches Wesen über grenzenlose kognitive Fähigkeiten und Selbstkontrolle verfügt, sodass die Rekonstruktion der eigentlichen Präferenzen der Individuen unmöglich gemacht wird. Dies gilt nicht nur für potentielle Entscheidungsarchitekten, sondern auch für die betreffenden

---

24 Vgl. Sugden (2008: 232), (2009: 370), Grüne-Yanoff (2012: 642f.), McQuillin/Sugden (2012: 560f.), Qizilbash (2012: 654) und White (2013: 76–8).

Individuen selbst; selbst die Betroffenen können ihre eigenen wahren, in dieser Art idealisierten Präferenzen nicht eruieren. Selbst wenn man die Anforderungen an rationale, informierte Präferenzen abschwächte, bliebe das grundlegende Problem, dass der Entscheidungsarchitekt die wahren Interessen der verschiedenen Individuen nicht kennen kann (vgl. White 2013: 61ff.). Um diese These zu erläutern, ist eine genauere Betrachtung des in Abschnitt 3 eingeführten Präferenzkonzepts der modernen Ökonomik angezeigt.<sup>25</sup>

Ein Individuum, so wurde dargelegt, handelt rational, wenn es gemäß einer stabilen und konsistenten Präferenzordnung entscheidet. Damit nun die Konsistenzanforderung dabei nicht leer wird und damit sich Präferenzen, wie es die ökonomische Theorie annimmt, tatsächlich in Entscheidungen offenbaren, müssen sie in einem umfassenden Sinne als Präferenzen „all things considered“ verstanden und über Alternativen definiert werden, die alle entscheidungsrelevanten Faktoren beinhalten (vgl. Hausman 2012: 15ff.). Diese Forderung sei anhand zweier Beispiele veranschaulicht. Erstens sei angenommen, eine Person stünde vor der Aufgabe, eine Präferenzordnung über die Eissorten Vanille, Schokolade und Erdbeere zu bilden. Nun wäre es möglich, in Abhängigkeit von verschiedenen Kriterien ganz unterschiedliche Rangordnungen zu bilden. Hinsichtlich des Geschmacks zieht sie vielleicht Vanille gegenüber Schokolade und Schokolade gegenüber Erdbeere vor, in Bezug auf die Konsistenz jedoch präferiert sie Schokolade gegenüber Vanille und Vanille gegenüber Erdbeere. Betrachtet die Person wiederum den Kaloriengehalt, so könnte sich die letztgenannte Präferenzordnung ins Gegenteil verkehren. Trifft das Individuum nun bei mehrmaligem Eiskonsum unterschiedliche Ent-

---

25 Siehe dazu auch Klonschinski (2016: 46ff.).

scheidungen auf der Basis verschiedener Ordnungsprinzipien, kann das Transitivitätsaxiom offenbar nicht mehr greifen. Die Präferenzordnung muss also unter Berücksichtigung aller Kriterien zugleich gebildet werden. Zweitens könnte eine Person ihre Präferenzordnung grundsätzlich einzig und allein ihren egoistischen Eigeninteressen gemäß bilden. Wenn sie jedoch in einer konkreten Situation, beispielsweise aus Gründen des Pflichtgefühls, nicht die in dieser Ordnung obenstehende Alternative wählt, dann offenbart sie ihre – zu eng definierten – Präferenzen nicht, sodass sich aus der Beobachtung dieser Entscheidung auch nicht auf die Präferenzordnung schließen lässt.

Diese Beispiele zeigen, dass, damit die Konsistenzanforderungen sinnvoll angewendet werden und Präferenzen ihre Funktion im Rahmen der ökonomischen Theorie erfüllen können, Präferenzen als „total comparative evaluations“ verstanden werden müssen (ebd.: 4). Die Alternativen, über die die Präferenzen gebildet werden, müssen dabei alle Faktoren umfassen, die das Individuum für seine Entscheidungen als relevant erachtet (vgl. ebd.: 18f.). Daraus folgt zum einen, dass die Konsistenzanforderungen an Präferenzen auf einer sehr allgemeinen Ebene gültig sind. Zum anderen ergibt sich das sowohl für die Verhaltensökonomik im Allgemeinen als auch für den LP im Speziellen zentrale Problem, dass der außenstehende Beobachter nicht weiß, welche Komponenten der Entscheidungssituation das Individuum als relevant erachtet und wie es die Alternativen separiert.<sup>26</sup> Entscheidungen und Präferenzen können damit nicht einfach beobachtet werden. Sagoff (2004: 65f.) illustriert diese These wie folgt: in Boston ließ sich im Jahr 1966 ein plötzlicher Einbruch der Nachfrage nach Fisch beobachten. Wie sich herausstellte, resultierte dieser Einbruch aus der Tatsache, dass

---

26 Vgl. Broome (1991a: 95ff.), Sagoff (2004: 58ff.), Mandler (1999: 89f.) und Pattanaik (2013: 217).

der Papst im selben Jahr das Verbot, freitags Fleisch zu essen, aufgehoben hatte. Die Entscheidungsalternativen stellten sich damit für die (katholischen) Individuen folgendermaßen dar: 1. Fleischessen mit päpstlichem Einverständnis, 2. Fischessen mit päpstlichem Einverständnis und 3. Fleischessen ohne päpstliches Einverständnis. Der zentrale Punkt ist, dass ohne Kenntnis der individuellen Verpflichtung dem Papst gegenüber die getroffene Entscheidung nicht beobachtet werden kann (vgl. ebd.: 73). Sowohl Entscheidungen als auch Präferenzen erweisen sich damit als etwas zutiefst Subjektives, da letztlich nur das Individuum selbst weiß, wie es die Alternativen definiert bzw. welche Gründe und Motive seinen Entscheidungen zugrunde liegen. Ihre „Beobachtung“ erfordert daher ein hohes Maß an Interpretation.

Hier ließe sich einwenden, dass die Verhaltensökonomik gerade diesem Problem Abhilfe schafft, indem sie, wie oben dargelegt, die kognitiven Prozesse der Individuen fokussiert. Da die Verhaltensökonomik jedoch zugleich dem behavioristischen Erbe der Ökonomik verhaftet bleibt (vgl. Güth/Kliemt 2013), bleibt sie mit der Analyse der Gründe hinter den Entscheidungen auf halbem Wege stehen, wie auch Grüne-Yanoff (2012: 637) anmerkt: „behavioral science rarely offers an account of the cognitive mechanisms underlying [...] behavioral regularities. It is not clear what makes discount hyperbolically or value losses more than gains.“ Daher ist auch nicht klar, wann eine bestimmte Entscheidung auf irrationalen Verzerrungen beruht und wann sie einer reflektierten, rationalen Überlegung entspringt. Eine geringe Sparrate für die Rente, um ein Beispiel von White (2013: 70, 100) aufzugreifen, kann Ausdruck von Irrationalität, aber auch Ergebnis einer rationalen Kalkulation sein, etwa, wenn das Individuum plant, sich in Kürze eine Eigentumswohnung zu kaufen.

Das bisher Gesagte zusammenfassend ist festzustellen, dass Präferenzen generell nicht beobachtbar sind und sie in ihrer von T&S vertretenen, hochidealisierten Variante noch nicht einmal das Individuum selbst kennen kann. Auch die Verhaltensökonomik kann keinen Aufschluss über die tatsächlichen Handlungsgründe der Individuen liefern. Um nichtsdestotrotz Aussagen über individuell „bessere“ Entscheidungen machen zu können, müssen T&S folglich eine substantielle Wohlergehenkonzeption vertreten. Und tatsächlich tun sie das auch, wenn sie darauf verweisen, dass sie den Individuen zu einem *längeren* und *gesünderem* Leben verhelfen wollen (vgl. Qizilbash 2012: 656). Die Postulierung dieser Ziele mutet auf den ersten Blick unkontrovers an – es scheint plausibel, jedem ein Interesse an einem langen Leben bei guter Gesundheit zu unterstellen. Bei näherem Hinsehen tun sich indessen zwei Probleme auf. Zum einen ist nicht jeder Weg, die Realisierung dieser Ziele zu erreichen, für jedes Individuum gleich gut. Während Spinat etwa im Allgemeinen ein gesundes Lebensmittel ist, kann er für Einzelne durchaus schädlich sein (vgl. Coons/Weber 2013: 18). Zum anderen lassen T&S außer Acht, dass Individuen auch noch andere Ziele haben, die miteinander in Konflikt geraten können.<sup>27</sup> So konfligiert das Interesse an Gesundheit und einem langen Leben bspw. mit dem Interesse an Lebensfreude oder sozialer Teilhabe; hier sei auf Risikosportarten oder unterhaltensreiche lange Nächte an der Bar verwiesen. Isoliert betrachtet und unter Voraussetzung der *Ceteris-paribus*-Bedingung mögen die genannten Ziele also plausibel sein, „all things considered“ sind sie es nicht. Interessanterweise schreiben T&S (2009: 2) im Kontext des Cafeteria-Beispiels selbst, dass es darum gehe, die Schüler durch die Anordnung der Lebensmittel „all things

---

27 So auch White (2013: 69, 73, 100).

considered“ möglichst gut zu stellen, beziehen sich in ihren Beispielen jedoch jeweils nur auf isolierte Ziele.

Die von T&S postulierten Interessen der Individuen spiegeln also letztlich eine bestimmte externe Vorstellung davon wider, was gut für die Betroffenen ist. Tatsächlich ist eine solche Setzung bestimmter Interessen unumgänglich, solange sich T&S am Rationalitätskonzept der modernen Ökonomik orientieren.<sup>28</sup> Wie dargelegt fordert dieses Konzept lediglich die Konsistenz von Präferenzen, sagt über deren Inhalt aber nichts aus. Individuen können präferieren, was sie wollen, auch wenn es „schlecht für sie“ ist, solange sie es konsistent tun (vgl. Hausman 2012: 19, 42). Ohne die Annahme konkreter Ziele, d.h. allein auf Basis der formalen Rationalitätstheorie, ist es nun aber unmöglich, Verhalten zu erklären oder zu prognostizieren (vgl. Davis 2003: 31). Will die Ökonomik dennoch Erklärungen liefern, so muss sie zusätzliche Annahmen über Ziele und Motive der Individuen einführen. Meist löst sie dieses Problem durch den Rekurs auf materielle Eigeninteressen (vgl. Hausman/McPherson 2008: 236; White 2013: 65f.); T&S „lösen“ es, wenn auch unbewusst, indem sie den Individuen implizit bestimmte, nur auf den ersten Blick überzeugende Ziele unterstellen. Sie füllen damit die Leerstelle, die die zunehmende Formalisierung der Nutzen- und Rationalitätskonzeption gelassen hat.

## 6. Fazit

Abschließend gilt es, die wesentlichen Ergebnisse kurz zusammenzufassen und ihre Relevanz für die Debatte um den LP herauszustellen. Der Beitrag hat die Entwicklung der Nutzen-

---

28 White (2013: 74) vertritt eine ähnliche These, untermauert diese aber nicht mit dogmengeschichtlichen Überlegungen.

theorie bzw. des Rationalitätskonzepts in der positiven und normativen Ökonomik vom Ende des neunzehnten Jahrhunderts bis heute skizziert. Diese Entwicklung war von dem Bestreben gekennzeichnet, die ökonomische Analyse weitestgehend von psychologischen Annahmen zu befreien. Im positiven Strang der Disziplin wurde aus einer hedonistischen Nutzentheorie, die dezidiert dazu gedient hatte, das Verhalten der Konsumenten zu erklären, eine formale Theorie, die Entscheidungen repräsentiert. Über Gründe und Motive des Handelns sagt diese formale Theorie nichts mehr aus. Die Verhaltensökonomik strebt in gewisser Weise eine Revision dieser Entwicklung an, indem sie tatsächliche Entscheidungen in den Blick nimmt und zu ihrer Erklärung auf psychologische Determinanten von Verhalten recurriert. Die Rationalitätskonzeption der Mainstream-Ökonomik, so wurde herausgearbeitet, bleibt dabei jedoch als normatives Ideal erhalten – erst dadurch können Abweichungen von diesem Standard als „Fehler“ charakterisiert werden. Da die moderne Wohlfahrtsökonomik Wohlergehen als die Erfüllung konsistenter Präferenzen definiert, führen derartige Abweichungen zum sogenannten Vereinbarkeitsproblem von Verhaltens- und Wohlfahrtsökonomik, auf welches der LP reagiert. Durch die entsprechende Entscheidungsarchitektur, so die Idee, ließen sich die Individuen zu rationalen, ihren eigentlichen konsistenten Präferenzen entsprechenden Entscheidungen bewegen. Abgesehen davon, dass die Erfüllung konsistenter Präferenzen nicht unbedingt gut für das Individuum sein muss, kann der Entscheidungsarchitekt individuelle Präferenzen jedoch weder beobachten, noch kann er die „wahren“ Präferenzen kennen – und in der von T&S vertretenen Variante können dies noch nicht einmal die Individuen selbst. T&S müssen somit eine substantielle Wohlergehenskonzeption voraussetzen, wenn sie Aussagen über „bessere“ Entscheidungen treffen wol-

len. Dies resultiert aus ihrer Orientierung am Rationalitätsideal der ENT: Da dieses keine Aussagen über Ziele und Gründe der Individuen erlaubt, müssen diesbezüglich zusätzliche Annahmen eingeführt werden.

Der LP lässt sich folglich ebenso wenig wie andere Formen des Paternalismus mit Rekurs auf die wahren Präferenzen der Betroffenen legitimieren und erfüllt damit den von T&S formulierten Anspruch nicht. Es sei jedoch darauf hingewiesen, dass die vorgebrachten Argumente keine Implikationen für die Legitimierbarkeit alternativer, auf substantiellen Wohlergehenskonzeptionen beruhender Varianten des Paternalismus haben. Um der hier vertretenen Kritik zu entgehen, müsste sich der LP in der von T&S vertretenen Form folglich von der Rationalitätskonzeption der ENT lösen und die implizit unterstellte Wohlergehenskonzeption explizieren und zur Diskussion stellen. Auf dieser Basis wäre dann seine Rechtfertigbarkeit unter Rückgriff auf die Frage, inwiefern der LP freiheitserhaltend ist, zu diskutieren.

### *Literatur*

- Berg, N.; Gigerenzer, G. (2010): As-if behavioral economics: neo-classical economics in disguise? *History of Economic Ideas* 17(1): 133–165.
- Blaug, M. (1997): *Economic Theory in Retrospect*, 5. Auflage. Cambridge, Cambridge University Press.
- Broome, J. (1991a): *Weighing Goods: Equality, Uncertainty and Time*. Oxford: Blackwell.
- (1991b): Utility. *Economics and Philosophy* 7: 1–12.
- Bruni, L., Sugden, R. (2007): The road not taken: how psychology was removed from economics, and how it might be brought back. *The Economic Journal* 117: 146–173.

- Camerer, C.F.; Loewenstein, G. (2004): Behavioral economics: past, present, future. In: Camerer, C.F.; Loewenstein, G.; Rabin, M. (Hg.): *Advances in Behavioral Economics*. Oxford: Oxford University Press, S. 3–51.
- Coons, C.; Weber, M. (2013): Introduction: paternalism – issues and trends. In: Dies. (Hg.): *Paternalism. Theory and Practice*. Cambridge: Cambridge University Press, S. 1–24.
- Dams, J.; Ettl, A.; Greive, M.; Zschäpitz, H. (2015): Merkel will die Deutschen durch Nudging erziehen. *Die Welt* vom 12.03.2015, online: [www.welt.de/wirtschaft/article138326984/Merkel-will-die-Deutschen-durch-Nudging-erziehen.html](http://www.welt.de/wirtschaft/article138326984/Merkel-will-die-Deutschen-durch-Nudging-erziehen.html), Zugriff am 20.11.2015.
- Davis, J. (2003): *The Theory of the Individual in Economics: Identity and Value*. London/New York: Routledge.
- Edgeworth, F. Y. (1881): *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*. London: Kegan Paul.
- Ellsberg, D. (1954): Classic and current notions of „measurable“ utility. *Economic Journal* 64: 528–556.
- Giocoli, N. (2003): *Modeling Rational Agents: From Interwar Economics to Early Modern Game Theory*. Cheltenham/Northampton: Edward Elgar.
- Grüne-Yanoff, T. (2012): Old wine in new casks: libertarian paternalism still violates liberal principles. *Social Choice and Welfare* 38: 635–645.
- Güth, W.; Kliemt, H. (2013): Behaviorism, optimization and policy advice. Unpublished Manuscript of a Talk held at the Radein Workshop 2013.
- Hausman, D.M. (2012): *Preference, Value, Choice and Welfare*. Cambridge: Cambridge University Press.
- Hausman, D.; McPherson, M. (2008) The philosophical foundations of mainstream normative economics. In: Hausman, D. (Hg.): *The Philosophy of Economics: An Anthology*, 3. Auflage. Cambridge: Cambridge University Press, S. 226–269.
- (2009): Preference satisfaction and welfare economics. *Economics and Philosophy* 25: 1–25.
- Hausman, D.M.; Welch, B. (2010): Debate: To nudge or not to nudge. *The Journal of Political Philosophy* 18(1): 123–136.

- Heukelom, Floris (2014): *Behavioral Economics: A History*. Cambridge: Cambridge University Press.
- Hicks, J.R. (1939): The foundations of welfare economics. *The Economic Journal* 49: 696–712.
- (1946): *Value and Capital: An Inquiry into some Fundamental Principles of Economic Theory*, 2. Auflage (1. Auflage 1939). Oxford: Clarendon Press.
- Jevons, W.S. (1911/2006): *The Theory of Political Economy*, Reproduktion der 4. Auflage (1. Auflage 1871). London: Macmillan.
- Kahneman, D. (2011): *Thinking, Fast and Slow*. London: Allen Lane.
- Kaldor, N. (1939): Welfare propositions of economics and interpersonal comparisons of utility. *The Economic Journal* 49: 549–552.
- Klonschinski, A. (2013): Das Kosten-Nutzen Verhältnis als Priorisierungskriterium? Eine philosophisch-dogmenhistorische Betrachtung des Ökonomischen Paradigmas der Wertmaximierung. In: Schmitz-Luhn, Björn; Bohmeier, André (Hg.): *Priorisierung in der Medizin: Kriterien im Dialog*. Berlin: Springer, S. 79–107.
- (2016): *The Economics of Resource Allocation in Health Care: Cost-Utility, Social Value, and Fairness*. Advances in Social Economics Series. Hg. von John B. Davis. London; New York: Routledge.
- Maas, H. (2005): *William Stanley Jevons and the Making of Modern Economics*. Cambridge : Cambridge University Press.
- Mandler, M. (1999): *Dilemmas in Economic Theory: Persisting Foundational Problems of Microeconomics*. Oxford : Oxford University Press.
- McQuillin, B.; Sugden, R. (2012): Reconciling normative and behavioral economics: the problems to be solved. *Social Choice and Welfare* 38: 553–567.
- Pareto, V (1927/1971): *Manual of Political Economy*, Übersetzung der Französischen Ausgabe von 1927 (1. Auflage 1906). London/Basingstoke: Macmillan.
- Pattanaik, P.K. (2013): The concepts of choice and preference in economics. *Journal of Economic Methodology* 10(2): 215–218.
- Pigou, A.C. (1932/1962): *The Economics of Welfare*, Nachdruck der 4. Auflage (1. Auflage 1920). London: Macmillan.

- Plickert, P.; Beck, H. (2014): Kanzlerin such Verhaltensforscher. In: *Frankfurter Allgemeine* vom 26.08.2014, online: [www.faz.net/aktuell/wirtschaft/wirtschaftspolitik/kanzlerin-angela-merkel-sucht-verhaltensforscher-13118345.html](http://www.faz.net/aktuell/wirtschaft/wirtschaftspolitik/kanzlerin-angela-merkel-sucht-verhaltensforscher-13118345.html), Zugriff am 19.11.2015.
- Qizilbash, M. (2012): Informed desire and the ambitions of libertarianism. *Social Choice and Welfare* 38: 647–658.
- Reinhardt, U. (1992): Reflections on the meaning of efficiency: can efficiency be separated from equity? *Yale Law and Policy Review* 10: 302–315.
- Robbins, L. (1935) *Essay on the Nature and Significance of Economic Science*, 2. Auflage. London: Macmillan.
- Ross, D. (2014): *Philosophy of Economics*. New York: Palgrave Macmillan.
- Sagoff, M. (2004): *Price, Principle, and the Environment*. Cambridge: Cambridge University Press.
- Samuelson, P.A. (1938): A note on the pure theory of consumer's behaviour. *Economica* 5(17): 61–71.
- (1947): *Foundations of Economic Analysis*. Cambridge: Harvard University Press.
- (1948/1968): Consumption theory in terms of revealed preferences. In: Page, A.N. (Hg.) (1968): *Utility Theory: A Book of Readings*. New York et al.: John Wiley & Sons, S. 149–157.
- Schabas, M. (1990): *A World Ruled by Number: William Stanley Jevons and the Rise of Mathematical Economics*. Oxford: Princeton University Press.
- (2007): *The Natural Origins of Economics*. Chicago/London: University of Chicago Press.
- Selinger, E.; Whyte, K. (2012): What counts as a nudge? *American Journal of Bioethics* 12(2): 11–31.
- Sugden, R. (2008): Why incoherent preferences do not justify paternalism. *Constitutional Political Economy* 19: 226–248.
- (2009): On nudging: a review of „nudge: Improving Decisions about Health, Wealth, and Happiness“ by Richard H. Thaler and Cass R. Sunstein. *International Journal of the Economics of Business* 16(3): 365–373.

Sunstein, C.; Thaler, R. (2003): Libertarian paternalism is not an oxymoron. *The University of Chicago Law Review* 70: 1166–1187

Thaler, R.; Sunstein, C.R. (2003): Libertarian paternalism. *American Economic Review Papers and Proceedings* 93(2): 175–179.

--- (2009): *Nudge. Improving Decisions about Health, Wealth, and Happiness*. London et al.: Penguin Books.

Von Neumann, J., Morgenstern, O. (1953) *Theory of Games and Economic Behavior*, 3. Auflage. (1. Auflage 1944). Princeton: Princeton University Press.

White, M.D. (2013): *The Manipulation of Choice. Ethics and Libertarian Paternalism*. New York: Palgrave.