

RESEARCH

Lena Dreher

Neuroökonomie

Eine wissenschaftstheoretische
Analyse

OPEN ACCESS



Springer Gabler

Neuroökonomie

Lena Dreher

Neuroökonomie

Eine wissenschaftstheoretische
Analyse

 Springer Gabler

Lena Dreher
Universität Konstanz
Konstanz, Deutschland

Dissertation der Universität Konstanz,

1. Referent: Prof. Dr. Wolfgang Spohn, 2. Referent: Prof. Dr. Urs Fischbacher, 3.
Referentin: Prof. Dr. Dorothea Debus,
Tag der mündlichen Prüfung: 21. September 2021



Förderung durch Manfred-Ulmer-Stipendium der Stiftung „Wissenschaft und Gesellschaft an der Universität Konstanz“, ein Stipendium im Rahmen der Begabtenförderung der Promotionsförderung der Konrad Adenauer Stiftung sowie Abschlussunterstützung des Ausschuss für Forschungsfragen (AFF) der Universität Konstanz

ISBN 978-3-658-37802-8 ISBN 978-3-658-37803-5 (eBook)
<https://doi.org/10.1007/978-3-658-37803-5>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

© Der/die Herausgeber bzw. der/die Autor(en) 2022. Dieses Buch ist eine Open-Access-Publikation. **Open Access** Dieses Buch wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Buch enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten. Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geographische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung/Lektorat: Marija Kojic
Springer Gabler ist ein Imprint der eingetragenen Gesellschaft Springer Fachmedien Wiesbaden GmbH und ist ein Teil von Springer Nature.
Die Anschrift der Gesellschaft ist: Abraham-Lincoln-Str. 46, 65189 Wiesbaden, Germany

Danksagung

Ich danke meinen Betreuern Prof. Wolfgang Spohn und Prof. Urs Fischbacher, die in diesem interdisziplinären Thema nicht nur fachlich herausragend kompetent sind, sondern auch bei jeder Frage oder Sorge mit Rat zur Seite standen. Eine bessere Betreuung hätte man sich nicht wünschen können.

Für die finanzielle Förderung danke ich der Stiftung „Wissenschaft und Gesellschaft an der Universität Konstanz“ für das Manfred-Ulmer-Stipendium, der Promotionsförderung der Konrad Adenauer Stiftung für das Stipendium im Rahmen der Begabtenförderung und dem Ausschuss für Forschungsfragen (AFF) der Universität Konstanz für die Abschlussunterstützung.

Das in dieser Arbeit verwendete generische Maskulinum dient ausschließlich der besseren Lesbarkeit und umfasst ausdrücklich weibliche und männliche sowie nicht-binäre Personen.

Lena Dreher

Inhaltsverzeichnis

1	Einleitung	1
2	Warum Neuroökonomie? Die Ausgangssituation in den Mutterwissenschaften	7
2.1	Warum sich Ökonomen für Neuroökonomie interessieren	7
2.2	Warum sich Neurowissenschaftler für Neuroökonomie interessieren	15
2.3	Der neuroökonomische Lösungsansatz	18
3	Verhaltensökonomie	25
3.1	Simon und die <i>bounded Rationality</i>	26
3.2	Smith und Laborexperimente	27
3.3	Kahneman und Tversky, Heuristiken und Verzerrungen	30
3.4	Selten und die Spieltheorie	41
3.5	Soziale Präferenzen	46
4	Neurowissenschaften	53
4.1	Fragen und Historie der kognitiven Neurowissenschaften	54
4.2	Ziele der kognitiven Neurowissenschaften	65
4.3	Empirische Methoden	81
5	Die neuroökonomische Zusammenarbeit	93
5.1	Der Beitrag der Verhaltensökonomie zur Neuroökonomie	94
5.2	Der Beitrag der kognitiven Neurowissenschaften zur Neuroökonomie	102
5.3	Wie nutzen Verhaltensökonomie und kognitive Neurowissenschaften die Vorteile des jeweils anderen?	122

6 Kritiken an der Neuroökonomie	129
6.1 Die prozeduralen Kritiken	131
6.2 Die systematischen Kritiken	151
6.3 Wissenschaftstheoretische Kritiken an der Neuroökonomie	167
7 Schluss	187
Literaturverzeichnis	197



Einleitung

1

Seit dem Jahr 2010 steigt die Zahl der Publikationen zum Thema kognitive Neurowissenschaften stark an, wie beispielsweise eine Recherche auf PubMed Central, dem Repositorium der US National Library of Medicine, zeigt. Mit ihnen angestiegen ist auch die Zahl der Veröffentlichungen aus dem Bereich der Neuroökonomie, einer noch relativ jungen Disziplin, die wir in dieser Arbeit näher betrachten werden.

Diese Arbeit ist eine wissenschaftstheoretische Analyse der Neuroökonomie. Sie betreibt selbst keine neuroökonomische Forschung, sondern analysiert die Neuroökonomie als Wissenschaft. Ich gehe dabei unter anderem den Fragen nach: Was ist Neuroökonomie? Nach welcher Art Modell suchen die Wissenschaftler und mit welchen Methoden gehen sie dabei vor? Wie funktioniert der Erkenntnisgewinn? Außerdem untersuche ich, welchen Kritiken die Neuroökonomie ausgesetzt ist und wie diese Kritiken zu beurteilen sind. Wie sind die Ziele und Vorgehensweisen der Neuroökonomie wissenschaftstheoretisch einzuordnen? Und nicht zuletzt stellt sich die Frage, welche Entwicklung und Relevanz angesichts steigender Publikationszahlen rund um kognitive Neurowissenschaften in Zukunft von der Neuroökonomie zu erwarten sind.

Die Neuroökonomie ist noch vergleichsweise unbekannt, aber insofern besonders, als dass die Personen, die an ihr forschen, aus verschiedenen Mutterwissenschaften stammen und noch immer Teil dieser Wissenschaften sind – daher halte ich es für passend, von einer neuroökonomischen Zusammenarbeit zu sprechen. Es ist nicht ganz einfach, abzugrenzen, welche Mutterwissenschaften die Neuroökonomie hat und welche Forscher aus welchen Gebieten an ihr tätig sind. Es sind sicherlich die Neurowissenschaften, genauer gesagt, die kognitiven Neurowissenschaften und die Ökonomie, insbesondere die Verhaltensökonomie. Wir werden in Kapitel 6 sehen, dass diese Spezifizierung essentiell wichtig für die Neuroökonomie ist. Vermutlich könnte man auch die kognitive Psychologie zu

den besteuernden Wissenschaften zählen. Für eine Auseinandersetzung mit drei Mutterwissenschaften wäre in dieser Arbeit allerdings kein Platz gewesen, sodass ich mich im Folgenden auf die kognitiven Neurowissenschaften und die Verhaltensökonomie als Mutterwissenschaften der Neuroökonomie konzentriere, aus welchen sich auch der Begriff Neuroökonomie zusammensetzt. Beide Wissenschaften, beziehungsweise ihre Forschenden, bringen unterschiedliche Methoden, Erfahrungen und Erkenntnisse mit in die Kooperation ein, ebenso wie unterschiedliche Fragen, die sie mithilfe neuroökonomischer Forschung klären oder zumindest näher beleuchten möchten. Das klingt einerseits nach einer Umgebung für kreative Forschung, birgt aber andererseits auch Konfliktpotenzial, wenn die Hintergründe oder wissenschaftlichen Methoden des jeweils anderen nicht verstanden werden. Wir werden in Kapitel 5 besprechen, inwiefern das durchaus zu Problemen führen kann.

Ganz generell ist der unterschiedliche Hintergrund der beteiligten Forscher eine prägende Charakteristik der Neuroökonomie. Daher werden wir in den folgenden Kapiteln die Neuroökonomie häufig in ihre Mutterwissenschaften aufteilen. Das heißt, wir werden die Verhaltensökonomie und die kognitiven Neurowissenschaften zunächst einzeln betrachten, um dann die Zusammenarbeit in der Neuroökonomie – und auch die Kritiken daran – besser zu verstehen.

Zunächst beginnen wir unsere Analyse der Neuroökonomie mit der Frage, wofür die Neuroökonomie steht, was sie erreichen möchte, was ihre Ziele sind. Dazu betrachten wir in Abschnitt 2.3 verschiedene Aussagen von Neuroökonominnen sowie ihre Einschätzung der Leistungsfähigkeit der Neuroökonomie, auf die wir weiter unten in Kapitel 6 zu den Kritiken an der Neuroökonomie zurückkommen werden. Um die Ziele der Neuroökonomie besser zu verstehen, vor allem auch, weshalb gerade diese Ziele den Wissenschaftlern wichtig sind, ist es sinnvoll, sich zuerst mit den Ausgangssituationen der Ökonomen und Neurowissenschaftler in ihren Mutterwissenschaften zu befassen. Daher werden wir in Abschnitt 2.1 und 2.2 einen kurzen Blick darauf werfen, welche Motivation Ökonomen beziehungsweise Neurowissenschaftler haben, sich mit neuroökonomischer Forschung zu beschäftigen. In der Ökonomie sind das vor allem eine seit Jahrzehnten vorhandene Unzufriedenheit mit der perfekten Rationalität des Modells *Homo oeconomicus* und die Suche nach Modellen, die näher am Verhalten realer Menschen liegen. Die Neuroökonomie bietet Ökonomen biologische Daten zum Verhalten realer Individuen. Die kognitiven Neurowissenschaften auf der anderen Seite haben seit der Erfindung nicht-invasiver Messmethoden eine Reihe an Möglichkeiten zur Verfügung, Daten über die Arbeit des menschlichen Gehirns zu generieren, benötigen für ihre Forschung aber einen passenden theoretischen Überbau wie Experimentparadigmen und Analysemethoden. Beide Seiten

hoffen, die neuroökonomische Zusammenarbeit könnte ihnen helfen, ihre Fragen zu beantworten.

Ab Kapitel 3 geht es dann vor allem um die Frage, wie die Mutterwissenschaften arbeiten und welche Voraussetzungen, Methoden und Erfahrungen sie in die Neuroökonomie einbringen. Das Kapitel 3 behandelt diese Fragen für die Verhaltensökonomie, das Kapitel 4 für die kognitiven Neurowissenschaften. Den vier Aspekten von Edward Cartwright (Cartwright 2014) folgend, werden wir sehen, wie sich die Verhaltensökonomie entwickelt hat: erstens über die Forderung von Herbert Simon nach einer *bounded Rationality* als Alternative zur perfekten Rationalität des *Homo oeconomicus*. Simon sah keine Beweise dafür vorliegen, dass in einer realen Entscheidung tatsächlich die geforderten Berechnungen durchgeführt werden oder überhaupt durchgeführt werden können. Zweitens über die Etablierung von Laborexperimenten vor allem durch Vernon Smith, der begann, ökonomische Hypothesen und Vorhersagen mithilfe systematischer empirischer Forschung zu überprüfen. Drittens über die Arbeit von Daniel Kahneman und Amos Tversky, die empirische Belege dafür lieferten, dass echte Menschen systematisch die Anforderungen an rationales Entscheiden nicht erfüllten und daher nicht denken und entscheiden wie *Homo oeconomicus*, und viertens über die entscheidenden Entwicklungen in der Spieltheorie durch John Nash, Reinhard Selten und John Harsanyi. Sie entwickelten die Spieltheorie zu einem nützlichen und vielseitigen Werkzeug für theoretische und empirische ökonomische Forschung. Dieser Blick auf die Entwicklung der Verhaltensökonomie hilft uns, Aufschluss darüber zu erhalten, welchen Fragen die Verhaltensökonomie heute warum nachgeht und welche Expertise die Ökonomen in die neurowissenschaftliche Zusammenarbeit einbringen.

Kapitel 4 befasst sich als Gegenstück dazu mit Entwicklung und Arbeit der kognitiven Neurowissenschaften. Auch hier wird die Historie der Wissenschaft dabei helfen, heutige Fragestellungen zu verstehen. Im ersten Teil des Kapitels geht es daher um die Fragestellungen der kognitiven Neurowissenschaften und wie sie sich aus den ersten Forschungen im 19. Jahrhundert über Phrenologie zu Konnektionismus und Behaviorismus entwickelt haben. Das schafft die Grundlage für den zweiten Teil des Kapitels zur gegenwärtigen Arbeit der kognitiven Neurowissenschaften, das sich um die Frage dreht, welche Art von Modell kognitive Neurowissenschaftler suchen und wie ihre Fragestellungen dazu aussehen. Das wird durch die Beispiele der kontemporären Forschung zu Gedächtnis und Emotionen illustriert, die auch für die neuroökonomische Forschung von Bedeutung sind. In diesem Kontext werden auch die beiden großen Projekte beziehungsweise Initiativen „Human Brain Project“ und „BRAIN Initiative“ vorgestellt. Der dritte Teil des Kapitels ist ein Überblick über heute in den kognitiven

Neurowissenschaften gängige technische Instrumente. Ich habe diejenigen Techniken ausgewählt, die auch in neuroökonomischer Forschung häufig eingesetzt werden: funktionelle Magnetresonanztomografie (fMRT), Positronenemissionstomografie (PET), Elektroenzephalografie (EEG) und transkranielle magnetische Stimulation (TMS) als nicht-invasive Methoden sowie die Ableitung von Aktionspotenzialen als invasive Messtechnik. Dieser Überblick lohnt sich nicht nur zum generellen Verständnis des Ablaufs kognitiver Experimente, sondern auch für das Verständnis mancher Kritiken, die auf dieser Technik beruhen. Diese Kritiken werden wir in Kapitel 6 besprechen.

Zuvor zeigt Kapitel 5 nach der Einzelbetrachtungen der Mutterwissenschaften die gemeinsame neuroökonomische Zusammenarbeit. Abschnitt 5.1 zeigt den Beitrag der Verhaltensökonomie zur Neuroökonomie. Wir werden sehen, inwiefern sich moderne ökonomische Experimente von den Experimenten Seltens und Smiths aus Kapitel 3 unterscheiden und wie mit klaren Benchmarks Emotionen operationalisiert werden. Dazu dient uns ein Experiment von Fehr zum Thema Fairness als Beispiel. Demgegenüber gestellt ist der *Balloon Analogue Risk Task* aus der Psychologie als Beispiel für Konstrukt-basiertes Experimentieren. Abschnitt 5.2 zeigt den Beitrag der kognitiven Neurowissenschaften zur Neuroökonomie. Dieser besteht vor allem aus der Generierung und Bereitstellung von Experimentdaten. Daher behandelt dieses Unterkapitel vor allem Schritte und Anforderungen bei kognitiven Experimentdesigns, beispielsweise wie Stimuli präsentiert werden oder wie Versuchsdurchläufe generell aussehen. Dabei konzentrieren wir uns auf das fMRT, da es die Messtechnik ist, mit der in einem beispielhaften neuroökonomischen Experiment gearbeitet wurde, das wir in Abschnitt 5.3 besprechen. Dieses Beispiel soll uns als Veranschaulichung dienen, wie Neurowissenschaftler ein Experiment aufbauen und mit ökonomischer Methode umsetzen und welche Schlüsse sich daraus ziehen lassen.

Nachdem wir gesehen haben, was die Neuroökonomie ist und wie sie arbeitet, widmet sich Kapitel 6 den Kritiken an der Neuroökonomie. Wir werden eine Reihe an Kritiken, ohne Anspruch auf Vollständigkeit, in prozedurale und systematische Kritiken ordnen und analysieren. Worauf fußen die Kritiken, sind sie überzogen oder berechtigt und was könnten Neuroökonom tun, um ihnen zu begegnen? Die prozeduralen Kritiken in Abschnitt 6.1 sprechen nicht die systematischen Grundlagen der Neuroökonomie an und können grundsätzlich gelöst werden. Dabei geht es beispielsweise um Kritiken an überzogenen Versprechungen vonseiten mancher Neuroökonom oder um neurowissenschaftliche Forschungsstandards, wie die angemessene Zahl an Probanden oder eine unzureichende technische Ausrüstung, die die Experimentergebnisse nicht-reproduzierbar macht. Die grundlegenden, systematischen Kritiken werden in Abschnitt 6.2

besprochen. Sie liegen weniger in der täglichen wissenschaftlichen Arbeit begründet, sondern in Eigenschaften der Mutterwissenschaften. Wir stützen uns dabei vor allem auf bekannt gewordene Kritiken der Ökonomen Faruk Gul und Wolfgang Pesendorfer, die beklagen, dass Neuroökonomien ökonomische Modelle und Theorien missverstehen und nicht hilfreiche Änderungen am gängigen ökonomischen Modell vorschlagen. Sie argumentieren für die Irrelevanz neurowissenschaftlicher Erkenntnisse für die Ökonomie, da die Ökonomie keine Aussagen über Gehirnfunktionen macht und daher auch nicht mit Gehirndaten widerlegt werden kann. Diese Kritiken lassen sich nicht aus dem Weg räumen, da sie, wie ich argumentieren werde, auf Gul und Pesendorfers behavioristischer Konzeption der Ökonomie (die ich Standardökonomie nenne) beruhen und mit der neuroökonomischen Konzeption der Ökonomie unvereinbar ist. Doch selbst wenn man mit den Verhaltensökonomien übereinstimmt und die Standardökonomie ablehnt, bleiben noch immer wissenschaftstheoretische Fragen und Probleme für eine Zusammenarbeit von Ökonomie und Neurowissenschaften bestehen. Diese Probleme werden in der Wissenschaftstheorie tatsächlich seit Jahrzehnten immer wieder diskutiert, mit Argumenten, die auch in diesem Fall gültig sind. Diese werden in Abschnitt 6.3 behandelt. Dabei geht es um Reduktionismus und seine Probleme Emergenz, multiple Realisierbarkeit, die Unmöglichkeit von Brückengesetzen und nicht zuletzt um Inkommensurabilität. Daher betrachten wir zunächst, wie sich Neuroökonom Paul Glimcher eine Zusammenarbeit vorstellt und wie er zu reduktionistischen Problemen steht, ehe wir uns unter anderem Argumenten von Kenneth Schaffner, Paul Feyerabend, Thomas Kuhn und Jerry Fodor zuwenden, der bereits im vorigen Jahrhundert die Frage diskutierte, ob neurowissenschaftliche Erkenntnisse dazu verwendet werden können, psychologische Modelle zu entwickeln.

Das Kapitel 7 wird als Schluss einen Ausblick auf die künftige Entwicklung der Neuroökonomie geben. Es wird sicherlich Kritiken und Probleme geben, die weiterhin bestehen und die Zusammenarbeit beeinflussen. Dazu gehören auch die systematischen und manche der prozeduralen Kritiken aus Kapitel 6. An welchen Problemen müssen die Wissenschaftler arbeiten und wie wird sich ihr Verhältnis in Zukunft entwickeln? Sicher ist, dass Themen rund um die kognitiven Neurowissenschaften weiterhin große Aufmerksamkeit bekommen werden.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Warum Neuroökonomie? Die Ausgangssituation in den Mutterwissenschaften

2

Um zu verstehen, welche Ziele die Neuroökonomie verfolgt und warum, ist es hilfreich, zu Beginn einen Blick auf die Ausgangssituationen der beteiligten Wissenschaften zu werfen. In diesem Kapitel werden wir genau das tun und dabei der Frage nachgehen, warum sich Verhaltensökonominnen und kognitive Neurowissenschaftler für die Neuroökonomie interessieren und welche Möglichkeiten sie sich von ihr versprechen. Dabei werden wir zum besseren Überblick getrennt nach den beiden Wissenschaften vorgehen, beginnend mit den Neurowissenschaften in Abschnitt 2.1, gefolgt von der Ökonomie in Abschnitt 2.2. Abschnitt 2.3 wird dann aufzeigen, welche Ziele sich Neuroökonominnen stecken und welche Lösungsansätze sie dabei verfolgen.

2.1 Warum sich Ökonomen für Neuroökonomie interessieren

Einfach ausgedrückt, beschäftigen sich Ökonomen mit Neuroökonomie, weil sie sich davon neue Erkenntnisse für die Ökonomie versprechen. Vor allem möchten sie mit den Ergebnissen der Neuroökonomie Fragen aus der Verhaltensökonomie beantworten, für die das Wissen gerade der kognitiven Neurowissenschaftler hilfreich sein könnte. Es lohnt sich ein kurzer Blick in die Historie der Ökonomie, um zu sehen, welche Fragen das sind und wie sie zustande kamen.

In den 1930er Jahren wurde in der Ökonomie ein Meilenstein gesetzt mit der Entwicklung der *Revealed-Preference*-Axiome, vor allem durch den amerikanischen Ökonomen Paul Samuelson. Im Gegensatz zu vorherigen Modellen konnten nun mit relativ einfachen Axiomen klare Vorhersagen über das zu erwartende Handlungsmuster eines ökonomischen Akteurs getroffen werden. Auf dieser vielversprechenden Grundlage erfolgte der nächste große Schritt in den 1940er Jahren

mit der Ausarbeitung der *Expected-Utility-Theory* und der Grundlegung der Spieltheorie durch John von Neumann und Oskar Morgenstern. Unter den neuen Theorien müssen Akteure, die den zur Theorie gehörigen Axiomen unterliegen, ein Verhalten zeigen, als ob sie eine stetige Nutzenfunktion hätten und als ob ihre Handlungen auf die Maximierung des erzielten Gesamtnutzens ausgerichtet wären. (Glimcher et al. 2009b, 2). Mit diesen Errungenschaften auf der relativ unkomplexen Basis von ein paar Axiomen hatte die neoklassische Ökonomie Theorien zur Untersuchung und zur Vorhersage von Wahlverhalten ökonomischer Individuen an der Hand und war damit leistungsfähiger als je zuvor.

Allerdings bestand schon damals unter Ökonomen die Meinung, das Bild, das sich die Ökonomen von den handelnden Personen mache, sei nicht realistisch¹, basiere auf überholter Psychologie und müsse an neue psychologische Entwicklungen angepasst werden (Friedman 1953, 30). Ökonom Milton Friedman (Friedman 1953, 15) zeigt sich in seinem Aufsatz „The Methodology of Positive Economics“ eher genervt von der Debatte und stellt klar:

„[...] the relevant question to ask about the ‚assumptions‘ of a theory is not whether they are descriptively ‚realistic,‘ [sic!] for they never are, but whether they are sufficiently good approximations for the purpose in hand. And this question can be answered only by seeing whether the theory works, which means whether it yields sufficiently accurate predictions.“

Für die Geltung der ökonomischen Theorien ist es für Friedman nicht wichtig, ob die ihnen zugeschriebenen Präferenzen und Wahrscheinlichkeiten von Ergebnissen tatsächlich in den Köpfen der handelnden Individuen anwesend sind. Diese sind zuvorderst Teil der Theorie. Für ihn ist relevant, dass sich das Verhalten der Personen so beschreiben lässt, als ob sie anwesend wären. Von einer nötigen Anpassung an zeitgenössische Psychologie kann daher für Friedman keine Rede sein, im Gegenteil. Er bringt das Beispiel der *Maximization-of>Returns*-Hypothese vor und sagt: Die andauernde Verwendung und Akzeptanz der *Maximization-of>Returns*-Hypothese über einen langen Zeitraum hinweg ist starkes Zeugnis für ihren Wert, ebenso wie das Scheitern der Entwicklung und breiter Akzeptanz jeglicher kohärenter, selbst-konsistenter Alternative. (Friedman 1953, 23)

Solange das ökonomische Modell eines handelnden Individuums also zufriedenstellende Ergebnisse in den Vorhersagen liefert, gibt es für Friedman keinen Grund, das Modell zu verändern oder gänzlich zu ersetzen. Mit dieser Haltung

¹ Friedman nennt die Diskussion „perennial“ und zitiert Kritiken aus den Jahren 1898 und 1947 (Friedman 1953, 30).

hat er stark dazu beigetragen, dass trotz aller Kritik in der Neoklassischen Ökonomie auch über die 1950er Jahre hinaus das Menschenbild des sogenannten *Homo oeconomicus* Bestand hat.

Werfen wir an dieser Stelle einen Blick auf das ökonomische Menschenbild, der dessen wichtigste Eigenschaften kurz zusammenfassen soll. Wir orientieren uns dabei an Kirchgässner (Kirchgässner 2013), der ein relativ facettenreiches, wenig strenges Bild von *Homo oeconomicus* zeichnet: *Homo oeconomicus* ist ein Modell für individuelles Verhalten von ökonomischen Akteuren. Das Verhalten wird dabei als „rationales Handeln“ (Kirchgässner 2013, 13) verstanden. In Entscheidungssituationen werden die Handlungsmöglichkeiten von Akteuren durch allgemeine und spezifische Restriktionen beschränkt, wie beispielsweise das verfügbare Einkommen der Akteure, die Preise auf den Märkten, oder rechtliche Beschränkungen. Innerhalb dieser Restriktionen können die Akteure Entscheidungen treffen und Handlungen ausführen. Dabei müssen sie aus ihren verschiedenen Handlungsmöglichkeiten wählen, wobei sie sich an ihren persönlichen Präferenzen orientieren. Diese können von Individuum zu Individuum verschieden sein, denn sie sind eine Fixierung der Wertvorstellungen des Individuums (Kirchgässner 2013, 14). Das heißt, sie gelten generell und werden nicht vor jeder Entscheidungssituation neu generiert.

Anhand der Präferenzen werden die Handlungsalternativen und ihre Konsequenzen vom handelnden Individuum durch Wertzuweisung nach Kosten und Nutzen beurteilt. Das Individuum entscheidet sich im Modell für diejenige Handlungsalternative, die nach dieser Abwägung den größten Nutzen nach Abzug der Kosten verspricht. Dabei sind einzelne Ziele dem großen Ziel der Nutzenmaximierung untergeordnet. Alle Entscheidungen, alle Einzelziele sind Mittel, um dieses große, lebenslange Ziel zu erreichen. Bei seinen Entscheidungen kennt das Individuum nicht notwendigerweise alle potenziellen Handlungsmöglichkeiten und kann auch nicht alle Konsequenzen zutreffend vorhersehen. Laut Kirchgässner (Kirchgässner 2013, 13–14) hat das Individuum aber fast immer die Möglichkeit, sich gegen Kosten mehr Informationen über Handlungsmöglichkeiten und mögliche Konsequenzen zu verschaffen.

Im ökonomischen Modell gehen Akteure bei ihren Entscheidungen rational vor. Dazu gehören zwei Punkte: Die Rationalität der Präferenzen und die Rationalität der Wahl. Die Präferenzen sind rational, wenn sie vollständig und transitiv sind. Um vollständig zu sein, muss ein Set von Präferenzen die Eigenschaft erfüllen: $A > B \vee A \sim B \vee B > A$ ². Diese Eigenschaft fordert, dass das handelnde Individuum zu jeder potenziellen Handlungsmöglichkeit eine Einstellung hat und sie

² A, B, C, D stellen verschiedene Handlungsmöglichkeiten dar. „>“ notiert Präferenz, „~“ notiert Indifferenz, „≥“ notiert schwache Präferenz.

in seinem Präferenzsystem verorten kann. Von jeder Handlungsmöglichkeit kann das Individuum sagen, ob es sie gegenüber anderen Möglichkeiten präferiert, nicht präferiert oder ob es ihr gegenüber indifferent ist. Weitere Möglichkeiten, wie beispielsweise „weiß nicht“, gibt es nicht.

Die Eigenschaft der Transitivität fordert folgende Eigenschaft von den Präferenzen:

(I) $A \geq B \wedge B \geq C \rightarrow A \geq C$ Transitivität der schwachen Präferenz

Die Transitivitäten von strikter Präferenz und Indifferenz werden analog gefordert:

(II) $A > B \wedge B > C \rightarrow A > C$ Transitivität der strikten Präferenz

(III) $A \sim B \wedge B \sim C \rightarrow A \sim C$ Transitivität der Indifferenz

(II) drückt die strikten Eigenschaften von (I) aus, während (III) die symmetrischen Eigenschaften von (I) ausdrückt (Hansson et al. 2012).

Das heißt, wenn das Individuum beispielsweise einen Apfel gegenüber einer Orange (schwach) präferiert und eine Orange gegenüber einer Birne, so wird es auch den Apfel gegenüber einer Birne (schwach) präferieren. Gleiches gilt für eine Indifferenzbeziehung zwischen den Wahlmöglichkeiten. Ist das Individuum indifferent zwischen einem Apfel und einer Orange, sowie zwischen einer Orange und einer Birne, so ist es auch indifferent zwischen einem Apfel und einer Birne. Einige weitere Eigenschaften folgen aus der Transitivität, als eine wichtige sei an dieser Stelle die Azyklizität erwähnt, die fordert: Es gibt keine Folge „ A_1, \dots, A_n “ von Handlungsalternativen, sodass „ $A_1 > \dots > A_n > A_1$ “ (Hansson et al. 2012).

Der zweite Punkt, die Rationalität der Wahl ist dann erfüllt, wenn das Individuum keine verfügbare Handlungsalternative derjenigen vorzieht (das heißt gegenüber derjenigen präferiert), die es gewählt hat (Hansson et al. 2012). Das bedeutet, im ökonomischen Modell wählt das handelnde Individuum immer eine Handlungsalternative, die es gegenüber allen anderen Alternativen (mindestens schwach) präferiert hat. Umgekehrt bedeutet das, dass die vom Individuum gewählte Handlung auch diejenige war, die es allen anderen Alternativen gegenüber (mindestens schwach) präferiert hat. Wählt das Individuum dagegen eine Alternative, die nicht die am meisten präferierte, sondern eine, angenommen, zweitbeste Lösung war, obwohl die erstbeste zur Verfügung gestanden hätte, so handelt es nicht rational.

Um einiges schwächer definiert Kirchgässner die Rationalität der Entscheidung:

„Rationalität bedeutet in diesem Modell lediglich, dass das Individuum, wenn es seinen Intentionen folgt, prinzipiell in der Lage ist, gemäss seinem relativen Vorteil zu handeln, d.h. seinen Handlungsraum abzuschätzen und zu bewerten, um dann entsprechend zu handeln.“ (Kirchgässner 2013, 17)

Hier ist lediglich gefordert, dass das Individuum gemäß seinem relativen Vorteil handeln *kann*. Zudem ist nichts gefordert über den Grad der Präferenz der gewählten Handlung gegenüber den übrigen Alternativen. Offenbar kann hier auch eine weniger präferierte Alternative gewählt werden, solange sie zum relativen Vorteil des Individuums beiträgt. Das ist bei der erstgenannten Definition der Rationalität der Wahl zwar auch möglich, doch handelt das Individuum dann nicht mehr rational. Bei Kirchgässner dagegen scheint es keine Verletzung der Rationalität zu sein, nicht die am stärksten präferierte Handlungsalternative zu wählen.

Kirchgässner sieht neben der Rationalität der Entscheidung einen weiteren wichtigen Punkt für die Entscheidung eines Individuums im ökonomischen Modell: Die Eigenständigkeit der Entscheidung.

„Eigenständigkeit bedeutet, dass das Individuum entsprechend seinen eigenen Präferenzen (und nicht entsprechend den Präferenzen anderer) handelt.“ (Kirchgässner 2013, 16)

Homo oeconomicus handelt ausschließlich nach seinen eigenen Interessen und wird als egoistisch angesehen, was bedeutet, dass das Individuum bei der Wahl einer Handlungsalternative grundsätzlich nicht darauf achtet, wie sich die Konsequenzen seiner Wahl auf andere Personen auswirken. Das Individuum sieht vielmehr nur seinen eigenen Nutzen und nicht den Payoff eventueller Mit- oder Gegenspieler. Es wird diejenige Handlung wählen, die für es selbst den Payoff maximiert. Ob der Gegenspieler dabei weniger oder auch mehr Payoff erhält als es selbst, ist für das Individuum irrelevant. Kirchgässner weist in diesem Zusammenhang darauf hin, dass das Wort „eigeninteressiert“ eine passendere Beschreibung sei, da es weniger moralisch belegt und die genauere Übersetzung der im Englischen gebräuchlichen Terminologie „self-interest“ ist (Kirchgässner 2013, 47).

Dies bedeutet jedoch nicht, dass die Interessen von Mit- und Gegenspielern überhaupt keinen Einfluss auf das Wahlverhalten von *Homo oeconomicus* haben. Grundsätzlich werden die Interessen anderer dort berücksichtigt, wo sie

den Handlungsspielraum des handelnden Individuums beeinflussen (Kirchgässner 2013, 16). In mehreren Aspekten können sie daher eine große Rolle spielen. Ein Aspekt, der die Wahl als Ganzes und weniger die eigene Präferenzbildung betrifft, ist, dass es bei vielen Entscheidungen, die das Individuum zu treffen hat, für den eigenen Payoff wichtig ist, welche Entscheidung die Gegenspieler treffen werden. Um die Handlungsalternative mit dem größten Payoff zu finden, muss das Individuum versuchen, zu antizipieren, wie der Gegner entscheiden wird. Das hängt von den Präferenzen, den Interessen des Gegenspielers ab, weshalb sie für das Wahlverhalten von *Homo oeconomicus* mindestens indirekt eine große Rolle spielen. Der zweite Aspekt ist vom ersten völlig verschieden und betrifft die Ebene der Präferenzbildung der Spieler. In diesem Aspekt geht es um *Other-regarding-Preferences*: Die Interessen anderer finden sich bereits in den Präferenzen des handelnden Individuums, beispielsweise weil es sich bei den berücksichtigten Personen um gute Freunde handelt, an deren Nutzen dem Individuum gelegen ist, und die es daher bei seinen Handlungen berücksichtigt. Das ist eine relativ moderne Idee aus der Verhaltensökonomie, wir werden uns in Abschnitt 3.5 (S. 50–52) näher mit Modellen mit sozialen Präferenzen beschäftigen. Kirchgässner spricht noch einen weiteren Aspekt an, bei dem vor allem die Reaktionen anderer Menschen auf die Handlungen des Individuums bei der Wahl einer Handlungsalternative eine Rolle spielen. Bei manchen Entscheidungen können gesellschaftliche Restriktionen vorliegen, die nicht gleichzeitig auch juristischen Restriktionen unterliegen. Das handelnde Individuum könnte eine mögliche Handlungsalternative vollkommen legal umsetzen, hat aber dann mit sozialen Sanktionen zu rechnen, die als Kosten angerechnet werden. Als Beispiel nennt Kirchgässner einen Lehrer, der seinen Schülern erklärt, dass es wichtig ist, wählen zu gehen, da es notwendig ist, um die Demokratie zu erhalten, dann aber selbst nicht zur Wahl geht. Dieser Lehrer hätte zwar keine rechtlichen Sanktionen zu erwarten, jedoch hohe soziale Kosten, weil er bei seinen Schülern stark an Glaubwürdigkeit verlieren würde. (Kirchgässner 2013, 61) Ohne diese sozialen Sanktionen wären dem Lehrer bei der Entscheidung, zur Wahl zu gehen oder nicht, die Interessen der Schüler relativ gleichgültig gewesen, sie kommen in seinem Präferenzset nicht vor. Allerdings sind ihre Reaktionen (und vermutlich auch die ihrer Eltern) Teil der Konsequenzen der Handlungsalternativen „wählen gehen“ und „nicht wählen gehen“. Daher haben sie dennoch Einfluss auf seine Entscheidung.

Wie gesagt, hat vor allem die strikte Version des Modells *Homo oeconomicus* viel Kritik erfahren. Sie beruht vor allem auf empirischer Evidenz aus Laborexperimenten, während derer die Probanden ökonomische Spiele spielen. Ergebnisse

aus solchen Studien widersprechen regelmäßig den Vorhersagen der Standardökonomie mit ihrer Idee vom *Homo oeconomicus*. Wir werden auf diese Spiele und ihre Bedeutung für die Ökonomie noch in späteren Kapiteln zu sprechen kommen, daher betrachten wir kurz die wichtigsten unter ihnen: das Ultimatumspiel, das Diktatorspiel, Vertrauensspiele und Spiele zum Thema öffentliche Güter.

Ultimatumspiele sind, wie alle Spiele in dieser Aufzählung, sogenannte Fairness-Spiele. Sie werden dazu genutzt, Erkenntnisse darüber zu gewinnen, wie Probanden Fairness auffassen. In diesem 2-Spieler-Spiel gibt es einen *Proposer* (P) und einen *Responder* (R). P erhält einen bestimmten Geldbetrag, sagen wir 10 €. Diesen kann er mit R teilen, indem er ihm einen bestimmten Betrag seiner Wahl anbietet. R kann das Angebot entweder akzeptieren oder ablehnen. Im ersten Fall erhält R den angebotenen Betrag und P den Rest des Geldes. Lehnt R das Angebot jedoch ab, erhalten beide Spieler kein Geld. Die ökonomische Standardtheorie sagt voraus, dass R auch sehr kleine Beträge akzeptieren wird, denn auch sie sind geschenktes Geld, das er im Falle einer Ablehnung nicht erhält. Daher wäre es für P rational, den kleinstmöglichen Betrag (in unserem Beispielfall 1 €) anzubieten, R sollte diesen Betrag dann akzeptieren. Jedoch zeigen Studien, dass P-Spieler bisweilen 50% des Geldbetrags anboten und R-Spieler oft positive Angebote ablehnten. Generell werden in Studien 60–80% des Betrags angeboten und Angebote von weniger als 20% häufig abgelehnt. Dies traf auch auf Studien mit höheren Geldbeträgen zu, die mehrere Monatseinkommen der Spieler umfassten. (Wilkinson 2008, 333). Das widerspricht den Vorhersagen des Standardmodells und wird so gedeutet, dass R-Spieler niedrige Angebote als unfair erachten und bereit sind, die P-Spieler dafür zu bestrafen, auch wenn das mit Kosten für sie selbst verbunden ist.

Auch bei Diktatorspielen gibt es einen *Responder* und einen *Proposer*, der einen Geldbetrag erhält, den er aufteilen kann. Allerdings hat R in diesem Spiel keine Möglichkeit, das Angebot zu abzulehnen oder zu akzeptieren, stattdessen wird ihm ein Betrag von P zugewiesen. Anders als das Ultimatumspiel soll das Diktatorspiel einen Einblick darin geben, ob P-Spieler nur deshalb großzügige Angebote machen, weil sie die Ablehnung durch R fürchten oder weil sie tatsächlich altruistisch sind, wobei jedes positive Angebot als altruistisch angesehen wird. In Studien wurden Angebote des Diktators von 20–40% des Geldbetrags beobachtet, jedoch mit 80–96% Null-Angeboten in Experimenten, in denen das verteilbare Geld von den P-Spielern erst selbst erarbeitet werden musste. (Wilkinson 2008, 334). Positive Angebote scheinen daher wenig altruistisch und zumeist strategisch begründet.

Für Vertrauensspiele charakteristisch ist, dass sie eine Investition von einem *Truster* (I) in einen *Trustee* (T) beinhalten. Die Investition ist für I risikobehaftet,

weil T sie zurückzahlen oder selbst behalten kann. Der Anreiz zur Investition ist jedoch, dass I von einer Investition profitieren kann, wenn der Return on Investment ausreichend ist und T genug zurückzahlt. In einem typischen Spielaufbau erhält I den Betrag x , den er investieren oder behalten kann. Wenn I Betrag y investiert, wird der Betrag um eine Dividende r des Betrags auf $y + yr$ vergrößert. Daraufhin muss T entscheiden, welchen Betrag davon er behalten und welchen Betrag er an I zurückzahlen möchte. Diese Spielsituation ist im Grunde ein Diktatorspiel. Angenommen, T behält Betrag z für sich und zahlt $y(1 + r) - z$ an I zurück, dann ist z der Lohn für T und $(x - y) + y(1 + r) - z$ für I. Interessant an diesem Spiel ist, dass es dazu verwendet wird, das Vertrauen, das I aufbringt, sowie die Vertrauenswürdigkeit von T zu messen: Das Vertrauen wird als Höhe der Investition y angesehen, während Vertrauenswürdigkeit anhand des zurückgezählten Betrags $x - z + ry$ gemessen wird. (Wilkinson 2008, 335). Das ökonomische Standardmodell sagt in Vertrauensspiel-Situationen null Vertrauen voraus. I-Spieler werden nicht investieren, weil T-Spieler nie etwas zurückzahlen werden. In Studien investieren I-Spieler durchschnittlich die Hälfte des ihnen zur Verfügung stehenden Betrags, wobei die Varianz sehr hoch ist mit 16% Spielern, die den gesamten Betrag investieren und 6%, die nichts investieren. T-Spieler zahlen im Schnitt 95% der investierten Summe zurück, was mehr ist als in einem üblichen Diktatorspiel, wobei der Unterschied mit circa 10% nicht sehr groß ist. (Wilkinson 2008, 335–336).

Bei Spielen zu öffentlichen Gütern geht es um den Beitrag der Spieler zu und die Nutzung von öffentlichen Gütern. Das sind solche Güter, die den Prinzipien der Nicht-Ausschließbarkeit und Nicht-Rivalität im Konsum unterliegen. Bei diesen Gütern ist es kaum möglich, einzelne Nutzer vom Gebrauch auszuschließen und mehrere Personen können diese Güter ohne Schaden für andere verwenden. Klassische Beispiele sind Straßen, Straßenbeleuchtung oder die Arbeit der Polizei. Für ökonomische Spiele interessant sind vor allem halb-öffentliche Güter, bei denen die (übermäßige) Nutzung durch Individuen negative Auswirkungen auf andere hat, wie beispielsweise Fischgründe, deren übermäßige Befischung durch einzelne Fischer sie für alle ruiniert. Auf der anderen Seite muss für die Bereitstellung und Funktionalität von öffentlichen Gütern ein Beitrag geleistet werden, zum Beispiel über Steuergelder. Die dominante Strategie laut ökonomischer Standardtheorie wäre, nichts zu Bereitstellung, Erhalt etc. der Güter beizutragen und so viel wie gewünscht zu konsumieren. Daraus entsteht ein Trittbrettfahrer-Problem: Nur wenige Spieler werden freiwillig für öffentliche Güter bezahlen, während viele Spieler zwar den Konsum genießen, aber nichts beitragen werden und davon profitieren, dass andere bezahlen. Ein typisches Spiel kann wie folgt aussehen: N Spieler erhalten je eine Summe x_i , von der sie den Betrag y_i in

ein öffentliches Gut mit dem Wert m pro Einheit investieren können. Spieler i erhält einen Gewinn von $x_i - y_i + m \sum y_j / N$, wobei $\sum y_j$ für die Summe aller Beiträge aller anderen Spieler steht. Die Vorhersage des Standardmodells wäre, wie gesagt, nichts beizutragen, obwohl alle Spieler von der Kooperation aller Spieler profitieren würden. Evidenz aus Studien zeigt, dass Spieler generell aber ungefähr 50% ihrer Summe x_i beisteuern, jedoch gibt es starke Variationen, da die meisten Spieler entweder nichts oder den gesamten Betrag einbringen. Es ist offenbar schwer zu sagen, warum Spieler kooperieren, sei es aus Altruismus oder weil sie erwarten, dass auch die anderen Spieler kooperieren werden. Ebenso scheint nicht klar zu sein, weshalb Spieler kein Geld beisteuern, aus dem Ziel heraus, kurzfristig den eigenen Gewinn zu maximieren oder weil sie erwarten, dass die anderen Spieler Trittbrettfahrer sind und auch nichts beitragen werden. (Wilkinson 2008, 338–339).

Während einige Ökonomen wie Friedman solche Diskrepanzen sowie eine größere Realitätsnähe des Modells *Homo oeconomicus* für nicht relevant halten und das in ihren Augen bewährte, perfekt rationale Modell weiterhin verwenden möchten, sind andere Ökonomen mit dieser Situation unzufrieden. Sie sind durchaus von der Relevanz der Realitätsnähe des ökonomischen Menschenbildes überzeugt und erforschen daher, wie ein realistischeres Modell aussehen könnte. Dabei sehen manche von ihnen in der Neuroökonomie eine potenziell hilfreiche Erkenntnisquelle, die, auf biologischen Prozessen basierend, näher an real handelnden Personen angelehnt ist als das Modell der Standardökonomie. In Kapitel 3 werden wir genauer sehen, wie dieser Zweig der Ökonomie, die Verhaltensökonomie, entstanden ist und welche Kompetenzen und Erfahrungen Verhaltensökonominnen in die neuroökonomische Zusammenarbeit einbringen.

2.2 Warum sich Neurowissenschaftler für Neuroökonomie interessieren

Neben den Ökonomen versprechen sich auch die Neurowissenschaftler von der Neuroökonomie Antworten auf Fragen aus ihrer eigenen Disziplin. Paul Glimcher, Colin Camerer, Ernst Fehr und Russel Poldrack geben einen kurzen Überblick darüber, welche Fragestellungen das sind und wie sie historisch in den Neurowissenschaften entstanden (Glimcher et al. 2009b, 1–12):

Seit dem 19. Jahrhundert gab es in der neurowissenschaftlichen Verhaltensforschung zwei unterschiedliche Richtungen, die Neurologie und die Physiologie. In der Neurologie wurden standardmäßig Menschen oder auch Tiere mit Hirnläsionen untersucht. Die meisten dieser Menschen hatten sich Verletzungen

durch Unfälle oder bei Kriegskämpfen zugezogen, aber auch Patienten mit fortgeschrittenen Hirntumoren waren potenzielle Probanden. Den Versuchspersonen wurde eine Reihe verhaltensbezogener Aufgaben gestellt. Die Defizite der Probanden bei der Bewältigung der Aufgaben wurden mit ihren jeweiligen Läsionen in Verbindung gebracht, was für Rückschlüsse auf die Funktionen der betroffenen Hirnregionen herangezogen wurde. Andersherum wurden bestimmte Stellen in den Gehirnen der Probanden mit Elektroden an Batterien angeschlossen und mit Stromstößen versehen. Aus den beobachteten Reaktionen, wie beispielsweise Muskelzuckungen, wurden Hypothesen über die Funktion der manipulierten Hirnregion aufgestellt. Die meisten dieser Experimente untersuchten Schäden an sensorischen oder motorischen Systemen. Ein Grund dafür war, dass die Auswirkungen von Schäden in diesen Systemen in Experimenten relativ leicht zu beobachten und zu quantifizieren sind, ebenso wie die sensorischen und motorischen Stimuli einfach kontrollierbar sind. Beispielsweise sind die Muskelzuckungen an den Hinterläufen eines Hundes, dessen Hirn man mit Stromstößen versieht, einfach zu beobachten und in ihrer Stärke einzuordnen. Mentale Zustände sind dagegen komplizierter zu kontrollieren und zu beobachten.

Die zweite Richtung, die Physiologie, untersuchte vor allem den Zusammenhang zwischen biologischen Zuständen der Probanden und Geschehnissen in der Umwelt. Beispielsweise wird den Probanden ein Bild vorgehalten, das als Stimulus dient, während das Feuern von Aktionspotenzialen, Veränderungen im Blutfluss oder der Neurotransmitter gemessen werden. Die experimentellen Möglichkeiten waren damals allerdings noch limitiert, da lediglich invasive Messmethoden zur Verfügung standen, die die biologische Materie beschädigen. Daher wurden solche Experimente ausschließlich an Tieren wie Affen, Meerschweinchen, Ratten oder Hunden durchgeführt.

Um die Zusammenhänge zwischen Gehirnaktivitäten und Verhalten trotz der experimentaltechnischen Einschränkungen erforschen zu können, wurden von den 1960er bis in die 1980er Jahre Modelle aus der Psychologie in den Neurowissenschaften eingeführt. Mit ihrer Hilfe erfolgten erste Schritte bei der Untersuchung von mentalen Zuständen. Auf Dauer treten jedoch zwei Probleme auf: Auf der einen Seite stehen zu viele Modelle zur Verfügung. Für einzelne Experimentergebnisse kommen mehrere Modelle infrage, die darauf zutreffen könnten. Auf der anderen, physiologischen Seite, stehen nicht genügend Daten zur Verfügung, die die Auswahl der Modelle bestimmen könnten. Das Vorhaben, diese Probleme zu lösen, war Kern der kognitiv-neurowissenschaftlichen Revolution (Glimcher et al. 2009b, 5). So wird unter anderem von Green und Swets 1966 die *Signal Detection Theory* entwickelt. Sie ist die erste globale normative Theorie, die dafür entwickelt worden war, neuronale Aktivitäten von Probanden in direkte Verbindung

mit beobachtetem Verhalten zu bringen. Ein Meilenstein wurde 1989 erreicht, als es William Newsome und seinem Team in einem Experiment an Affen zum ersten Mal gelang, wirklich klar eine Korrelation zwischen neuronaler Aktivität und stochastischer Wahl aufzuzeigen. Diese Art von Korrelation wurde als *psychometric–neurometric Match* bekannt. Das Experiment und seine Ergebnisse sind von hoher Relevanz in der neurowissenschaftlichen Forschung, weil es den ersten erfolgreichen Versuch darstellt, Entscheidungen aus der Aktivität einzelner Neuronen vorherzusagen (Glimcher et al. 2009b, 6).

Nach diesen Erfolgen wurde die *Signal Detection Theory* auch in Experimenten an Hirnregionen angewandt, die für die Bewegungsgenerierung zuständig sind. Allerdings waren die Ergebnisse hier weniger zufriedenstellend für die Neurowissenschaftler. Offenbar war es nicht einfach, die Experimente an für Motorik zuständigen Hirnregionen an eine auf der *Signal Detection Theory* basierenden Analyse anzupassen. Diese Rückschläge weckten in der neurowissenschaftlichen Gemeinschaft Zweifel an der Eignung der *Signal Detection Theory* für die umfassende Erforschung von Entscheidungsfindung. Es wird befürchtet, sie sei nicht ausreichend generalisierbar, um alle Aspekte eines Entscheidungsfindungsprozesses abzudecken (Glimcher et al. 2009b, 6).

Auf Seiten der Datengenerierung wurden unterdessen Fortschritte erzielt. In den 1990er Jahren kamen neue Technologien auf zur Messung neuronaler Aktivitäten am lebenden, wachen Menschen, wie beispielsweise die Positronenemissionstomographie (PET) und die heute vor allem aus Krankenhäusern bekannte funktionelle Magnetresonanztomographie (fMRT). Der Zugang zu diesen Technologien eröffnete neue Möglichkeiten der Datengenerierung und erweiterte den Pool untersuchbarer Fragestellungen. Wie diese von kognitiven Neurowissenschaftlern genutzt werden, werden wir genauer in Kapitel 4 sehen. Im Vergleich zu den Methoden des frühen 20. Jahrhunderts stehen den Forschern nun Unmengen an Daten zur Verfügung. Doch aus dem neuerlichen Segen wird bald das Problem evident, dass die Neurowissenschaftler kein theoretisches Werkzeug besitzen, um mit den Daten umzugehen (Glimcher et al. 2009b, 6). Dazu kommt noch das Problem, dass die menschliche Entscheidungsfindung eine kognitive Tätigkeit höherer Ebene ist als die Aktivität einzelner Neuronen. Sie ist ein relativ neues Betätigungsfeld für die Neurowissenschaftler, auf dem sie bisher weder viel Erfahrung noch viele Erkenntnisse besitzen. Immerhin bieten die technischen Entwicklungen der vergangenen Jahrzehnte nun effektive Technologien für die Datengewinnung. Das Problem der unzulänglichen Versorgung mit Daten durch physiologische Experimente konnte somit weitgehend gelöst werden. Was den Neurowissenschaftlern jetzt noch fehlt, ist ein umfassender theoretischer

Rahmen, der vorgibt, wie die neuronalen Messungen in den höheren, kognitiven Kontext zu setzen sind. Dazu gehört ein passendes theoretisches Werkzeug, um mit den generierten Datenbergen aus den neuen Technologien umgehen zu können.

Newsome ist der Meinung, die *Signal Detection Theory* könnte diese Leistungen erbringen. Andere Neurowissenschaftler sind jedoch eher skeptisch, ob dieses mathematische Standardwerkzeug stark genug generalisiert werden kann, um umfassende kognitive Prozesse abbilden zu können. Diese Neurowissenschaftler suchen daher nach neuen Ansätzen, wovon eine die neuroökonomische Kooperation mit Ökonomen ist. Denn für ihr Ziel, die neurobiologische Materie, die das Wahlverhalten hervorbringt, in unterschiedlich komplizierten Entscheidungssituationen zu beschreiben, sahen die Neurowissenschaftler um Glimcher vor einigen Jahren die neoklassischen ökonomischen Theorien mit der *Revealed-Preference*-Idee als potenziell nützlich an und integrierten sie in die neurowissenschaftliche Literatur (Glimcher et al. 2009b, 7).

2.3 Der neuroökonomische Lösungsansatz

Aus den gerade besprochenen Ausgangssituationen heraus sehen manche Verhaltensökonominnen und kognitive Neurowissenschaftler die Neuroökonomie als potenziell fruchtbare Zusammenarbeit, die neue Erkenntnisse und Forschungsmöglichkeiten für ihre Mutterwissenschaften liefern soll. Die Neurowissenschaftler auf der einen Seite erhoffen sich von der neuroökonomischen Zusammenarbeit neue Einblicke in die Arbeitsweisen des menschlichen Gehirns beim Ablauf komplexer Entscheidungsprozesse. Die ökonomischen Vorgehensweisen und Experimente zur Simulation ebensolcher komplexer Entscheidungssituationen sollen dabei weiterhelfen. Die Ökonomen auf der anderen Seite erhoffen sich Alternativen zur *Revealed-Preference*-Idee, die Grundlage für das ökonomische Menschenbild des *Homo oeconomicus* ist. Sie sind überzeugt, mit heutigen neurowissenschaftlichen Methoden die behavioristische Black Box öffnen zu können. Diese Ergebnisse sollen keine Hoffnungen bleiben. Die Befürworter der Neuroökonomie sind sich sicher, sie eines Tages auch vorlegen zu können. Tatsächlich sind die Ziele noch sehr allgemein formuliert, weshalb wir in diesem Unterkapitel genauer betrachten, wie die konkreten Vorhaben der Neuroökonominnen aussehen und was man sich darunter vorstellen muss, wenn sie sagen, sie wollen die Black Box öffnen.

Bei der Lektüre neuroökonomischer Veröffentlichungen, sei es ältere Literatur oder jüngere, fällt auf, dass einige Autoren, was die Ziele der Wissenschaft

angeht, bei allgemeinen Formulierungen bleiben. So schreiben beispielsweise Martin Reimann und Bernd Weber:

„Ziel der Neuroökonomie ist es, mittels neurowissenschaftlicher Methoden die so genannte behavioristische Black Box zu ‚öffnen‘ und affektive sowie kognitive Vorgänge im menschlichen Nervensystem besser zu verstehen.“ (Reimann et al. 2011, 7)

Auch Ökonom Douglas Bernheim wird wenig konkret, wenn er sagt, Neuroökonomie strebe danach, die Prozessstrukturen aufzudecken, die neuronale Aktivitäten determinieren, ebenso wie die neuronalen Prozesse, die Entscheidungen bestimmen (Bernheim 2009, 117). Natürlich ist nicht jede Publikation, die sich mit Neuroökonomie beschäftigt, an Fachleute aus den Mutterwissenschaften gerichtet, bei welchen ein gewisses Maß an Fachwissen vorausgesetzt werden kann, sondern auch an interessierte Laien, die lediglich einen Überblick bekommen möchten. Dennoch sind solche Allgemeinplätze enttäuschend nichtssagend und eine verpasste Chance, die Neuroökonomie nicht als neues Prestige-Projekt der Neurowissenschaften vorzustellen, sondern als unverzichtbare Wissenschaft mit Relevanz für die Zukunft. Vor allem dann, wenn die Lektüre in wissenschaftlichen Fachzeitschriften erscheint oder als Lehrbuch für Universitäten verkauft wird.

Manche Autoren werden konkreter in ihrer Beschreibung der neuroökonomischen Ziele. Die verschiedenen Aussagen stimmen zwar nicht genau überein, aus ihren Gemeinsamkeiten jedoch lässt sich ein Bild davon rekonstruieren, welche Ziele die Neuroökonomien konkret verfolgen. Dabei helfen Aussagen, die nicht nur die Ziele direkt betreffen, sondern vor allem den Nutzen, den die Neuroökonomie haben soll. Aussagen darüber, was sich die Forscher von der Neuroökonomie im Ganzen oder bestimmen Experimenten im Einzelnen versprechen, finden sich häufiger in der Literatur als Angaben darüber, was die Neuroökonomie als Wissenschaft erreichen möchte. So sind einige Ziele der Neuroökonomie über den Nutzen zu erkennen, den sich die Mutterwissenschaften erwarten.

Es lassen sich zwei große Kategorien öffnen: Den Nutzen, den sich die Neurowissenschaften versprechen und der Nutzen, der für die Ökonomie gesehen wird. Beide werde ich im Folgenden darlegen, beginnend mit dem Nutzen der Neuroökonomie für die Neurowissenschaften:

Auch hier gibt es wieder allgemein gehaltene Aussagen, die nicht viel Erkenntnis liefern. So sagt beispielsweise der Neurowissenschaftler Colin Camerer:

„The [neuroeconomic] approach should also inform neuroscience by expanding the range of cognitive activities that are studied.“ (Camerer 2005a, 12, Ergänzung von mir)

Der Zugewinn an kognitiven Aktivitäten für die neurowissenschaftliche Forschung mag anfangs wenig spektakulär klingen. Er wird für die Forscher aber interessant, nimmt man Camerers Aussage an anderer Stelle hinzu:

„Despite the rapid growth of game theory as an analytical tool at many social levels, we know almost nothing about how the human brain operates when people are thinking strategically in a game.“ (Camerer 2005a, 12)

In diese Richtung geht die oben erwähnte erweiterte Spanne der neurowissenschaftlich untersuchbaren kognitiven Aktivitäten. Es gibt für die kognitiven Neurowissenschaften noch viel Forschungspotenzial, wenn es um die Frage geht, wie das menschliche Gehirn in strategischen Entscheidungssituationen arbeitet. Ein großer Aspekt dieser Forschung wird sicherlich sein, menschliche Gehirne, die gerade mit einer strategischen Entscheidung und dem dazugehörigen Entscheidungsprozess beschäftigt sind, bei der Arbeit zu untersuchen, das heißt vor allem, die Neuronenaktivität zu messen. Dazu benötigen die Forscher zunächst einmal Experimentaufbauten mit spieltheoretischen Aufgaben für die Probanden. Das mag zunächst trivial klingen, kann die Forscher aber vor Schwierigkeiten stellen, vor allem dann, wenn sie noch nicht lange mit Spieltheorie arbeiten. Zumal die Anforderungen an ein erfolgreiches Experiment nicht allein mit den Aufgaben erfüllt sind, die die Probanden während des Experiments lösen sollen. Hinzu kommen die Theorien und Modelle, die hinter der spieltheoretischen Entscheidungsfindung stecken und die für den Aufbau, die Durchführung und die Analyse auch neurowissenschaftlicher Experimente verwendet werden müssen. Wer den theoretischen Hintergrund nicht kennt, kann ihn nicht berücksichtigen und wird somit bereits Schwierigkeiten haben, ein Experimentdesign festzulegen, das zu den verfolgten Forschungsfragen passt. Bei der Analyse der Daten sind Kenntnisse über die Theorie der Spieltheorie ebenso unerlässlich, da sonst die Frage offenbleibt, wie der Datensatz zu interpretieren ist, der mit Hilfe der Probanden generiert wurde. Kennen die Forscher beispielsweise nicht die verschiedenen Schritte, die die Spieltheoretiker einem Agenten bei der Entscheidungsfindung in spieltheoretischen Entscheidungssituationen zuschreiben, so werden sie ihre Experimente auch nicht auf Fragen nach einzelnen Schritten ausrichten können – und sie später natürlich auch nicht danach analysieren. Wenn die Forscher keine, oder eine minderwertige, Theorie zu den Abläufen parat haben,

die sie über die Daten legen können, dann haben sie wenig mehr als einen Stapel bunter Bilder aus dem fMRT, den sie nicht zu interpretieren wissen.

An wen sollten sich die Neurowissenschaftler also wenden, wenn nicht an diejenige Wissenschaft, die sich seit Jahrzehnten mit Spieltheorie in Theorie und Experiment befasst? Die Ökonomie liefert nicht nur den theoretischen Überbau, sondern auch erprobte Experimentdesigns, die sich auch in neurowissenschaftlichen Experimenten nutzen lassen. Camerer sieht große Hoffnungen aus dieser Kooperation erwachsen:

„And to many neuroscientists, the greatest promise of neuroeconomics is to supply theories and experimental designs for neuroscience. These neuroscientists feel that the kinds of models and tasks economists use routinely can contribute to ‘systems neuroscience’ understanding of higher-order cognition, which are challenging for neuroscientists who are used to focusing on very fine details of neurobiology and specific brain areas.“ (Camerer 2007, C26)

Camerer bestätigt hier auch, dass Neurowissenschaftlern die Arbeit mit kognitiven Tätigkeiten höherer Stufe schwerfällt, da sie es nicht gewohnt sind, sich mit ihnen zu beschäftigen. Die tägliche Arbeit der Neurowissenschaftler spielt sich üblicherweise auf einer detailreichen, tieferen Ebene ab, auf der der Fokus der Forscher auf der Tätigkeit einzelner Hirnbereiche oder einzelner Neuronen liegt. Ökonomen haben durch ihre eigene Arbeit eine globalere Sichtweise auf die Entscheidungsprozesse, die sie den Neurowissenschaftlern nahebringen könnten. Diese wiederum könnten dadurch einen langen Lernprozess verkürzen, da sie diese Sichtweise nicht erst selbst erarbeiten müssen.

Der Nutzen, den sich die Neurowissenschaftler von der Neuroökonomie erhoffen, liegt nach diesen Betrachtungen in der Anwendung ökonomischer Erkenntnisse, Methoden und Erfahrungen aus dem Bereich der Spieltheorie in den Neurowissenschaften. Die Neurowissenschaftler möchten in ihrer täglichen Arbeit vom Wissen der Ökonomen profitieren, um damit neuen Forschungsfragen nachgehen zu können. Das Ziel der Neuroökonomie in den Neurowissenschaften ist es, mit Hilfe ökonomischer Modelle und Theorien die Erkenntnisse über die Abläufe im menschlichen Gehirn in strategischen Entscheidungssituationen signifikant zu erweitern.

Auch die Ökonomen versprechen sich etwas von der Zusammenarbeit mit den Neurowissenschaftlern. Wie im vorangegangenen Kapitel besprochen, ist die Neuroökonomie auf der Suche nach einem neuen Menschenbild in der Ökonomie. Das neue Menschenbild soll, im Gegensatz zur Idee vom *Homo oeconomicus*, nahe an realen Menschen und ihren Verhaltensmustern angelegt sein. Schließlich

soll es auf reale Interaktionen zwischen realen Menschen anwendbar sein. Ökonomen könnten damit nicht nur theoretische Interaktionen zu Forschungszwecken untersuchen und die Ergebnisse präzise vorausberechnen. Auch wirtschaftliche Entscheidungen aus dem realen Leben könnten mit realistischem Menschenbild genauer beschrieben und sogar vorhergesagt werden, als mit dem abstrakten *Homo oeconomicus*. Beispiele für solche Entscheidungssituationen wären Fragen wie: *Wie reagieren meine Kunden/meine Mitbewerber, wenn ich den Preis meiner Produkte um X Euro erhöhe? Wird es die Arbeitskraft meiner Mitarbeiter erhöhen, wenn ich einen Betriebskindergarten einrichte?* Neue Modelle mit einem realistischen Menschenbild könnten nicht nur auf Kleingruppen oder Einzelpersonen, sondern auch auf ganze Märkte angewandt werden: *Wie werden die Märkte reagieren, wenn die Europäische Zentralbank den Leitzins erneut senkt (oder anhebt)? Werden geplante Reformen der Politik überhaupt die angestrebten Auswirkungen auf die Märkte haben oder werden sie ineffektiv verpuffen oder eine Situation gar verschärfen?* Solcherlei Fragen zutreffend beantworten zu können, hätte große Vorteile für Wirtschaft und Politik.

Was kann die Neuroökonomie tun, um präzisere Vorhersagen möglich zu machen? Dafür, sagt Glimcher, müssen die *As-if*-Modelle aus der Ökonomie abgelöst werden. Annahmen, Menschen handelten, *als ob* sie diese oder jene Berechnung in ihrem Kopf ausführten, sind nicht nahe genug am tatsächlichen Verhalten von Menschen. Statt eines künstlichen Gebildes, wie es der *Homo oeconomicus* ist, braucht es ein Modell, das so nah an realen Menschen ist, dass es ihr Verhalten akkurat beschreiben kann. Nur dann sind auch die akkuraten Vorhersagen möglich, nach denen in der Ökonomie gesucht werden. Glimcher ist sich sicher, dass die Neurowissenschaften diesen realistischen Einblick in das menschliche Verhalten liefern können, da sie das Know-How und die Technik besitzen, die behavioristische Black Box zu öffnen, die das menschliche Gehirn darstellt. Sie können dem Gehirn beinahe direkt bei der Arbeit zusehen und so Daten liefern, die seine Arbeitsweise direkt wiedergeben. Welches Menschenbild wäre näher am realen Menschen als eines, das direkt am realen Menschen abgenommen wurde? Zwar wurden auch die ökonomischen *As-if*-Modelle, die Glimcher hart kritisiert, aus der Beobachtung menschlichen Verhaltens gewonnen (*revealed Preferences*). Jedoch konnten dabei lediglich der Informationsinput und der Handlungoutput für das Erstellen des Modells berücksichtigt werden. Für den Prozess dazwischen, der sich im Gehirn der Handelnden abspielt, mussten Annäherungen verwendet werden. Genau das soll das neuroökonomische Modell so wertvoll machen. Vom Input bis zum Output soll es hier keine unbeleuchteten Stellen mehr geben, keine Black Box soll mehr ungeöffnet bleiben, an derer statt Annäherungen angestellt werden müssen. Die Neuroökonomie möchte darauf hinarbeiten,

dass es nicht mehr heißt „Proband X handelt so, *als ob* Prozess Y in seinem Kopf ablaufe“, sondern „Proband X handelt so, *weil* Prozess Z in seinem Kopf abläuft“. Oder, wie Glimcher es formuliert, er will die *As-if*-Modelle durch *Because*-Modelle ersetzen (Glimcher 2011, 130, 393). Die *Because*-Modelle beschreiben logische Operationen, die nicht nur Verhalten vorhersagen, sondern auch die darunterliegenden physischen und mentalen Prozesse genau nachempfinden, die dem Verhalten zugrunde liegen (Glimcher 2011, 393). Sie beschreiben also die Prozesse, die sich in der handelnden Person abspielen. Das deckt sich mit einer anderen Aussage Glimchers zum Ziel der Neuroökonomie:

„The goal of neuroeconomics is an algorithmic description of the human mechanism for choice.“ (Glimcher 2009c, 503)

Es lassen sich daraus grob zwei Ziele der Neuroökonomie herausarbeiten. Erstens die *Because*-Modelle, die logische Operationen beschreiben, die sich im Menschen abspielen, und zweitens eine algorithmische Beschreibung der Mechanismen, die sich im Menschen abspielen. Leider sagt bisher noch kein Neuroökonom etwas zu dem genauen Zusammenhang dieser beiden Ziele, evidenterweise lässt sich jedoch vermuten, dass die angestrebte algorithmische Beschreibung Grundlage für die *Because*-Modelle liefert, oder sie sogar darstellt. Denn im neurowissenschaftlichen Kontext wird es sich höchstwahrscheinlich sowohl bei den ‚logischen Operationen‘ als auch bei den ‚Mechanismen‘ um neuronale Aktivitäten im Gehirn handeln. Neuroökonomien möchten also die neuronalen Aktivitäten im menschlichen Gehirn beim Treffen von Entscheidungen algorithmisch beschreiben. Diese Beschreibung soll dann neue Modelle begründen, die in der Ökonomie zur Vorhersage menschlichen Verhaltens in Entscheidungssituationen verwendet werden.

Fasst man die Ziele beider Kategorien, die der Neurowissenschaften und die der Ökonomie, zusammen, so suchen die Neuroökonomien nach einer globalen Theorie über die Abläufe im menschlichen Gehirn beim Treffen strategischer Entscheidungen, die die neuroökonomischen Mutterwissenschaften und ihre Erklärungen für menschliches und tierisches Verhalten vereint (Glimcher 2011, 14–15). Laut Glimcher (Glimcher 2011, 393) ist das Ziel, ein einziges, einheitliches Modell menschlicher Entscheidungsfindung zu erstellen, das die ökonomische, psychologische und neurowissenschaftliche Analyse umfasst.

Es lässt sich also tatsächlich ein gemeinsames Ziel in der Neuroökonomie identifizieren, das die Erkenntnis sowohl in den Neurowissenschaften als auch in der Ökonomie mehren soll. Es wird nicht eindeutig gesagt, wie dieses Ziel erreicht werden soll. Doch es ist davon auszugehen, dass vor allem eine

immense Datengenerierung aus neurowissenschaftlichen Experimenten erfolgen muss, denn diese Daten sind die Grundlage für die algorithmische Beschreibung menschlichen Verhaltens in Entscheidungssituationen. Wie genau neuroökonomische Experimente aufgebaut sind und durchgeführt werden, werden wir in Kapitel 5 besprechen. Zuvor lohnt sich in den folgenden beiden Kapiteln ein Blick auf die Arbeit und Ziele der Mutterwissenschaften Verhaltensökonomie und kognitive Neurowissenschaften einzeln. Wie arbeiten sie, aus welchem wissenschaftlichen Hintergrund kommen sie und was können sie davon zur Neuroökonomie beitragen? Aus dem Verständnis dieser Hintergründe lassen sich die Arbeit und das Potenzial der neuroökonomischen Forschung besser einordnen.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





In Kapitel 2 wurden die Ziele der Neuroökonomie betrachtet, sowie die Erwartungen, die die Neurowissenschaftler und Ökonomen an die Neuroökonomie haben. Nun werfen wir einen Blick auf die Frage nach der Arbeit der Neuroökonomien für die Erreichung dieser Ziele. Dafür teilen wir die Neuroökonomie in ihre Mutterwissenschaften Ökonomie und Neurowissenschaften auf und betrachten die Arbeit der Neurowissenschaftler und Ökonomen zunächst eigenständig. In diesem Kapitel beginnen wir mit der Betrachtung der ökonomischen Seite der Neuroökonomie und sehen, wie die konkrete Arbeit der Ökonomen in dieser neuen Wissenschaft aussieht. In Kapitel 4 wird die Arbeit der kognitiven Neurowissenschaften thematisiert werden, worauf wir in Kapitel 5 beide Seiten zu einer Analyse der neuroökonomischen Zusammenarbeit zusammenbringen.

Doch zunächst zur Ökonomie: Um die Arbeit der Ökonomen besser verstehen zu können, ist es hilfreich, in Abschnitt 3.1 einen Blick auf die Verhaltensökonomie zu werfen, den Teilbereich der Ökonomie, der sich zumeist für die Neuroökonomie interessiert. Auf diese Weise erfahren wir, welches die ursprünglichen, ökonomischen Interessen und Kompetenzen dieser Wissenschaftler sind, die sie für ihre neuroökonomische Arbeit einsetzen. Für eine Charakterisierung der Verhaltensökonomie werden wir uns in Abschnitt 3.1 an vier Aspekten orientieren, die Edward Cartwright (Cartwright 2014) in einer Überblicksarbeit als Meilensteine in der Entwicklungsgeschichte der Verhaltensökonomie bezeichnet. Sie zeigen sehr gut die Probleme der Standardökonomie sowie die Lösungsansätze auf, die Verhaltensökonomien in den vergangenen Jahrzehnten vorgebracht haben. Das reicht von Herbert Simons Idee von *bounded Rationality* über die Etablierung von Laborexperimenten in der Ökonomie bis hin zu Daniel Kahnemans und Amos Tverskys *Heuristics-and-Biases*-Programm und der Entwicklung der Spieltheorie als wirksames Analyseinstrument.

Ein weiterer Aspekt moderner Verhaltensökonomie ist der Umgang mit vom Standardmodell abweichendem Verhalten von Versuchspersonen mittels der Entwicklung neuer Modelle zur Modellierung erweiterter Präferenzen im Vergleich zum Standardmodell. Dazu zählen beispielsweise soziale Präferenzen. In Abschnitt 3.2 werden zwei Modelle für die Einarbeitung sozialer Präferenzen in die Nutzenfunktionen ökonomischer Agenten vorgestellt. Sie zeigen, wie Verhaltensökonominnen mit Experimentergebnissen umgehen und welche Ideen neuen Modellen zugrunde liegen.

3.1 Simon und die *bounded Rationality*

In Abschnitt 2.2 betrachteten wir Probleme, die *Homo oeconomicus* als Menschenbild der Ökonomie mit sich bringt. Es wird als nicht realistisch angesehen und daher für wenig geeignet, Entscheidungssituationen aus der realen Wirtschaftswelt zu modellieren. Aus einer Frustration mit dieser Situation heraus entwickelte sich ab Mitte des 20. Jahrhunderts die Verhaltensökonomie.

Ökonom Edward Cartwright (Cartwright 2014) beschreibt die Entwicklungsgeschichte der Verhaltensökonomie anhand vier verschiedener, aber zusammenwirkender Aspekte, denen wir hier folgen werden, da sie das Geschehen recht anschaulich gliedern. Es geht dabei um die *bounded Rationality* von Simon, das *Heuristics-and-Biases*-Programm von Kahnemann und Tversky, die Etablierung von Laborexperimenten in der Ökonomie und die Entwicklungen, die die Spieltheorie im zwanzigsten Jahrhundert durchläuft, dank der Arbeiten von John Nash, Reinhard Selten und John Harsanyi. Beginnen wir mit dem ersten Aspekt:

Simon gilt als einer der ersten Kritiker des zeitgenössischen ökonomischen Menschenbildes. Er forderte, *Homo oeconomicus* durch ein realistischeres Modell zu ersetzen. Seiner Meinung nach gibt es keinen einzigen Beweis dafür, dass in einer tatsächlichen Entscheidungssituation die von den Ökonomen geforderten Berechnungen durchgeführt werden können oder durchgeführt werden. (Simon 1955, 104). Ein neues Modell sollte der eingeschränkten Rationalität realer Menschen Rechnung tragen, womit Simon den Begriff *bounded Rationality* prägte. Allerdings stieß Simons Aufruf unter den Ökonomen zunächst auf weitgehend taube Ohren, was seiner Idee damals den Einfluss versagte, den sie vielleicht verdient hätte. Das erklärt Cartwright damit, dass Simon keinen formalen Beweis dafür vorgebracht hatte, dass *Homo oeconomicus* keine gute Annäherung an das menschliche Verhalten ist. Das machte es für Ökonomen einfacher, Simons Arbeiten zu ignorieren. (Cartwright 2014, 7). Erst Kahneman und Tversky konnten

solche Belege liefern, mit einer Fülle an Experimenten und Hypothesen, die wir in Abschnitt 3.3 ausführlicher betrachten werden.

3.2 Smith und Laborexperimente

Der zweite Entwicklungsaspekt der Verhaltensökonomie, den Cartwright (Cartwright 2014) beschreibt, ist die Etablierung von Laborexperimenten als Werkzeug der empirischen Ökonomie. Maßgeblich war hier vor allem die Arbeit von Vernon Smith, für die er 2002 auch den Nobelpreis für Ökonomie erhielt. Smith war dabei gar nicht der erste Ökonom, der experimentelle Studien durchführte. Er selbst (Smith 1962) bezieht sich auf Edward Chamberlin, der Berichte über seine Experimente bereits in den Vierzigerjahren publiziert hatte (Chamberlin 1948). Allerdings weicht Smith bei seinem Experimentaufbau in einigen Punkten von dem Chamberlins ab, um Nachteile auszuschalten oder andere Marktaspekte besser beleuchten zu können. Smiths Studien begannen 1955 mit einer sechs Jahre dauernden Experimentreihe, die er darüber hinaus weiterzuentwickeln gedachte (Smith 1962, 111). Wir werfen hier einen kurzen Blick darauf, da Smiths Ideen und Vorgehensweisen wegbereitend für die Entwicklung der heutigen Experimentkultur in der Verhaltensökonomie waren:

Zum Teil sind diese Experimente darauf ausgerichtet, Hypothesen zum Marktgleichgewicht des Ökonomen Alfred Marshall zu testen, die dieser bereits 1890 veröffentlicht hatte (Marshall 1890). Da reale Märkte gewöhnlich sich verändernden Bedingungen unterworfen sind, definierte Marshall das Gleichgewicht als einen Zustand, zu dem sich der Markt hinbewegen würde, wenn Angebot und Nachfrage für eine ausreichend lange Zeitspanne unverändert blieben. Um testen zu können, ob sich ein Gleichgewicht tatsächlich mit der Zeit einstellen würde, hält Smith (Smith 1962, 115) in seinen experimentellen Märkten die Bedingungen von Angebot und Nachfrage über mehrere aufeinanderfolgende Handelsperioden konstant. Dafür teilt Smith, ähnlich wie Chamberlin, seine Versuchsteilnehmer nach dem Zufallsprinzip in zwei Gruppen auf, die Verkäufer und die Käufer. Jeder der Käufer erhält eine Karte mit einer Zahl darauf, die nur er kennt. Diese Zahl repräsentiert den maximalen Preis, den dieser Käufer gewillt ist, für eine Einheit der gehandelten Ware zu bezahlen. Den Käufern wird erklärt, dass sie keinen höheren Preis als den auf ihrer Karte für die Ware bezahlen dürfen, aber gerne einen niedrigeren Preis bezahlen möchten (je niedriger desto besser). Jedoch würden sie lieber genau den Preis auf der Karte bezahlen, als keine Ware zu erhalten und so ihre Bedürfnisse unbefriedigt zu lassen. Darüber hinaus wird den Käufern erklärt, dass sie sich vorstellen sollen, sie machen bei einem Kaufabschluss einen

Reingewinn in Höhe der Differenz zwischen dem Maximalpreis auf der Karte und dem tatsächlich erzielten Handelspreis. (Smith 1962, 112). Diese Erklärung sollte die Probanden dazu anhalten, nicht beim ersten Angebot kleiner oder gleich ihrem Maximalpreis zuzuschlagen, sondern auf einen möglichst niedrigen Preis hinzuwirken.

Analog erhalten die Verkäufer je eine Karte mit einer Zahl, die nur ihnen bekannt ist und den Minimalpreis darstellt, zu dem sie eine Einheit der gehandelten Ware zu verkaufen bereit sind. Auch ihnen wird erklärt, dass sie gewillt sein sollten, zu diesem Preis ihre Ware zu verkaufen, was ihnen lieber sein sollte, als am Ende der Handelsperiode keine Ware verkauft zu haben. Jedoch machen auch die Verkäufer einen Reingewinn, der der Differenz zwischen erzielttem Preis und Minimalpreis entspricht, sodass sie versuchen sollten, ihre Ware zu einem so hohen Preis wie möglich zu verkaufen. Unterhalb des Minimumpreises sollten sie aber auf keinen Fall einen Handel eingehen.

Für Smith (Smith 1962, 115) ist es wichtig, dass die Marktteilnehmer darüber hinaus keine Informationen über den fiktiven Markt erhalten. Es wird keine Historie angezeigt und niemand erfährt den Maximum- beziehungsweise Minimumpreis der übrigen Käufer beziehungsweise Verkäufer. Damit möchte Smith seine Experimentalmärkte nur mit solchen Informationen ausstatten, die auch in einem realen Markt praktisch erfahrbar wären. In einem realen Markt, beispielsweise einer Wertpapierbörse, kennen die Marktteilnehmer die Reservationspreise der übrigen Marktteilnehmer für gewöhnlich nicht. Um trotzdem Informationen über die Marktbedingungen zu erlangen, müssen sie die getätigten und akzeptierten Angebote und Gebote beobachten und so etwas über das Verhalten der anderen lernen. Aus diesen „public data“ (Smith 1962, 115) des Marktes sollen die Versuchsteilnehmer ihre Schlüsse ziehen.

Diese Bedingungen werden, wie gesagt, über mehrere Handelsperioden konstant gehalten. Die Käufer und Verkäufer behalten auch ihre Maximal- und Minimalpreis-Karten aus der ersten Periode für alle übrigen Perioden.

Im Unterschied zu Chamberlins Experimenten beschränkt Smith die Handlungsmöglichkeiten der Marktteilnehmer nicht auf bilaterale Gespräche, sondern ermöglicht multilaterales Bieten. Wenn der Markt für die erste Handelsperiode öffnet, dürfen die Marktteilnehmer ihre Hand heben und ein mündliches Angebot zum Kauf beziehungsweise Verkauf abgeben. Dabei ist jede Preisnennung erlaubt, die den eigenen Maximal- beziehungsweise Minimalpreis nicht verletzt. Die Marktteilnehmer dürfen nach eigenem Ermessen ein Angebot annehmen, wodurch ein Vertrag zwischen den beiden Handelspartnern abgeschlossen wird und sie den Markt für den Rest der Handelsperiode verlassen. Sobald ein Handel abgeschlossen ist, wird der Vertragspreis schriftlich festgehalten, sowie die

Minimal- und Maximalpreise der Vertragspartner. Dies wird weitergeführt, bis keine neuen Verträge mehr zustande kommen oder die Zeit einer Handelsperiode abgelaufen ist. In Smiths frühen Experimenten, die wir hier betrachten, besteht ein Experiment aus drei bis sechs Handelsperioden, abhängig davon, wie viele Versuchsteilnehmer anwesend sind. Da Smith die Experimente während einer Vorlesungssitzung durchführt, ist die Gesamtlaufzeit jedes Mal (gleich) beschränkt.

Im Rahmen dieser Arbeit werden wir nicht alle Resultate aus Smiths ersten Experimenten einzeln betrachten können, schauen aber auf fünf Hypothesen, die Smith (Smith 1962, 134) aus den Daten schloss: Erstens gibt es starke Tendenzen hin zu einem Marktgleichgewicht, sofern die Angebote und Gebote sowie die Transaktionen öffentlich gehalten und Preisabsprachen verhindert werden. Zweitens führen Veränderungen der Bedingungen von Angebot und Nachfrage zu entsprechenden Veränderungen in der Anzahl an Transaktionen und ihrem Preisniveau. Die neuen Preise entsprechen dabei recht genau den Vorhersagen der *Competitive Price Theory*. Drittens sieht Smith in den Daten leichte Belege dafür, dass für die Vorhersage eines statischen Gleichgewichts in einem Markt Wissen über die Form der Angebots- und Nachfragekurven und ihre Schnittpunkte voraussetzen. Viertens zeigen Märkte, in denen nur die Käufer Kursnotierungen vornehmen dürfen, schwächere Tendenzen zu einem Gleichgewicht, wenn nicht sogar Tendenzen dagegen, zugunsten der Käufer. Eine mögliche Erklärung für Smith ist, dass die Käufer in diesem Verfahren nur ein Minimum an Informationen über ihre Kaufbereitschaft preisgeben. Fünftens sieht Smith die von ihm so genannte *Walrasian Hypothesis*¹ nicht bestätigt. Er hält die *Excess Rent Hypothesis*² für adäquater.

Bei Smiths Experimenten und seinem Beitrag zum Aufstieg der Verhaltensökonomie ist eines auffällig: Im Gegensatz zu den Forschungen von Simon und Kahneman und Tversky legen seine Ergebnisse nahe, dass das Standardmodell der Ökonomie doch Einiges richtigmacht. Smith kommt nicht zu dem Schluss, man müsse das bisherige Modell grundlegend verändern, um realistische Vorhersagen treffen zu können. Aber er war einer der einflussreichsten Forscher, die Laborexperimente als empirisches Werkzeug in der Ökonomie etablierten. Seine Ideen, wie Experimente aufgebaut werden sollten, um bestimmte Markteigenschaften zu simulieren, gehören heute zum Standardrepertoire eines jeden experimentell

¹ Smith fasst sie zusammen: „[...] the rate of increase in exchange price is an increasing function of the excess demand at that price [...]“, Smith 1962, 116.

² Smith erläutert die *Excess-Rent-Hypothesis* so: „[It] [...] relates the ‘speed’ of contract price adjustment to the algebraic excess of buyer plus seller ‘virtual’ rent over the equilibrium buyer plus seller rent.“, Smith 1962, 134.

arbeitenden Ökonomen. Auch Smiths Ideen (Smith 1962, 120), die Motivation und Effizienz der Versuchsteilnehmer durch die Auszahlung von leistungsgebundenen monetären Belohnungen zu erhöhen, wird heute noch als Mittel der Wahl eingesetzt. Doch das ist nicht seine alleinige Errungenschaft für die Verhaltensökonomie: Smith begann nicht um ihrer selbst willen mit der Durchführung von Experimenten. Er entwickelte diese Idee, um bestehende ökonomische Hypothesen auf die Probe zu stellen und mögliche Alternativen zu identifizieren und erforschen. Das heißt, Smith nahm ökonomische Modelle als potenziell unrealistisch an und nahm an, sie müssten daraufhin untersucht werden. Es scheint ihm also wichtig gewesen zu sein, herauszufinden, ob die ökonomischen Modelle einen Realitätstest bestehen, und wie reale Menschen ökonomisch handeln. Das ist es, was ihn als Verhaltensökonom der frühen Stunde auszeichnet.

3.3 Kahneman und Tversky, Heuristiken und Verzerrungen

Den dritten der vier Aspekte in der Entwicklung der Verhaltensökonomie sieht Cartwright in der Arbeit der Psychologen Kahneman und Tversky. Sie zeigten vor allem auf, dass reale Menschen nicht wie *Homo oeconomicus* denken, rechnen oder entscheiden. Sie konzentrierten sich dabei vor allem auf Entscheidungen unter Unsicherheit und stützten ihre Aussagen auf Experimente, die sie oder Kollegen durchgeführt hatten. Mit diesen Experimenten zeigen sie, dass Menschen systematisch die Anforderungen an rationale Entscheidungsfindungen verletzen, auch Wissenschaftler, die im Umgang mit den gestellten Problemen geschult sind. Kahneman und Tversky (Tversky et al. 1974, 1130) stellten sich unter anderem gegen die Konzeption der subjektiven Wahrscheinlichkeit, die sie in der Entscheidungstheorie falsch angewandt sahen. Ihrer Meinung nach kommen die subjektiven Wahrscheinlichkeiten, die eine Person beispielsweise dem Eintreffen eines Ereignisses zuschreiben, dadurch zustande, dass sich diese Personen auf eine Anzahl an Heuristiken verlassen. Diese Heuristiken reduzieren die komplexe Arbeit der Wahrscheinlichkeitserfassung und Wertzuschreibung auf einfachere Entscheidungsaufgaben. (Tversky et al. 1974, 1124). Die Nutzung von Heuristiken bringt den Vorteil, komplexe Rechenaufgaben umgehen zu können, um sie durch Heuristiken zu ersetzen. Sie zu nutzen, ist nach Kahneman und Tversky (Tversky et al. 1974, 1131) höchst ökonomisch und für gewöhnlich auch effizient, führt aber auch zu systematischen und vorhersehbaren Fehlern, den kognitiven Verzerrungen (*cognitive Biases*). In zahlreichen Veröffentlichungen beschreiben Kahneman und Tversky eine Reihe von Heuristiken und auch

heute noch fügen vor allem Psychologen neue hinzu. Damit konnten Kahneman und Tversky auf eine große Datenbasis setzen, die ihre Thesen empirisch stützen konnten, wodurch sie einen besseren argumentativen Stand hatten als noch Simon. Da sie einen wichtigen Baustein in der Entwicklung der Verhaltensökonomie darstellen und um die Arbeit der beiden verstehen zu können, werden wir drei interessante Heuristiken und ihre Verzerrungen betrachten, die Kahneman und Tversky immer wieder in ihren Publikationen beschreiben: Die Repräsentativheuristik (*Representativeness Heuristic*), die Verfügbarkeitsheuristik (*Availability Heuristic*) und die Ankerheuristik (*Adjustment and Anchoring Heuristic*).

Die Repräsentativheuristik wird bei Fragen folgenden Typs verwendet: Mit welcher Wahrscheinlichkeit gehört Objekt A zur Klasse B? Mit welcher Wahrscheinlichkeit entstammt Ereignis A aus Prozess B? Wie hoch ist die Wahrscheinlichkeit, dass Prozess B Ereignis A hervorbringt? (Tversky et al. 1974, 1124). Kahneman und Tversky (Tversky et al. 1974, 1124) schreiben, dass in der Repräsentativheuristik Wahrscheinlichkeiten danach bewertet werden, in welchem Maße A repräsentativ für B ist, also in welchem Maße A B ähnelt. Wenn also A sehr repräsentativ für B ist, wird die Wahrscheinlichkeit, dass A aus B entstammt, als hoch eingeschätzt und umgekehrt. Ein Beispiel für eine Entscheidung mit der Repräsentativheuristik ist ein mittlerweile sehr bekanntes Experiment, das Kahneman und Tversky über die Jahre in verschiedenen Variationen und Kontexten durchgeführt haben.³ Eine Person wird von einem früheren Nachbarn wie folgt beschrieben (Tversky et al. 1974, 1124):

„Steve is very shy and withdrawn, invariably helpful, but with little interest in people or in the world of reality. A meek and tidy soul, he has a need for order and structure, and a passion for detail.“

Die Versuchsteilnehmer bekommen eine Beschreibung der obigen Art vorgelegt und sollen dann entscheiden, mit welcher Wahrscheinlichkeit Eigenschaften auf einer Liste auf die beschriebene Person zutreffen. In unserem Beispiel wurden sie gebeten, einzuschätzen, mit welcher Wahrscheinlichkeit Steve einem bestimmten Beruf nachgeht, beispielsweise Landwirt, Verkäufer, Flugzeugpilot bei einer Airline, Bibliothekar oder Arzt. Die Frage, die sich Kahneman und Tversky dabei stellen, ist, wie die Versuchsteilnehmer ein Ranking vornehmen, vom

³ Wir betrachten nur eine der verschiedenen Varianten, für weitere Experimente dieser Art siehe beispielsweise: Kahneman, D./Tversky, A. (1973) On the Psychology of Prediction. *Psychological Review* 80, 237–251; Tversky, A./Kahneman, D. (1982) Judgments of and by Representativeness. In: D. Kahneman et al., ed. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge etc.: Cambridge University Press, 84–98.

wahrscheinlichsten Beruf zum unwahrscheinlichsten. Unter Nutzung der Repräsentativheuristik, so sagen die beiden Psychologen, wird die Wahrscheinlichkeit, dass Steve beispielsweise ein Bibliothekar ist, danach beurteilt, wie ähnlich, oder repräsentativ er für das Stereotyp eines Bibliothekars ist. Sie fanden in weiteren Studien heraus, dass das Ranking der Berufe nach Wahrscheinlichkeit gleich ausfällt wie ein Ranking nach Ähnlichkeit. (Tversky et al. 1974, 1124). Dieses Verhalten zeigt auch gleich eine der Verzerrungen auf, die durch die Nutzung von Heuristiken auftreten können. Denn bei einem Ranking von Steves Berufen nach Wahrscheinlichkeit sollte die Anfangswahrscheinlichkeit dafür, dass Steve beispielsweise ein Bibliothekar ist, eine Rolle spielen. Die Tatsache, dass es in der Bevölkerung viel mehr Landwirte als Bibliothekare gibt, sollte in die Entscheidung, ob Steve eher ein Landwirt oder Bibliothekar ist, einfließen. Tatsächlich hat sie keinen Einfluss auf die Repräsentativheuristik, was zu Fehleinschätzungen führt.⁴ (Tversky et al. 1974, 1124). Neben der Anfangswahrscheinlichkeit (*Insensitivity to Prior Probability of Outcomes*) wird auch die Größe der Stichprobe offenbar nicht beachtet (*Insensitivity to Sample Size*). Kahnemans und Tverskys Experiment unter Bachelorstudenten macht das anschaulich (Tversky et al. 1974, 1125):

„A certain town is served by two hospitals. In the larger hospital about 45 babies are born each day, and in the smaller hospital about 15 babies are born each day. As you know, about 50 percent of all babies are boys. However, the exact percentage varies from day to day. Sometimes it may be higher than 50 percent, sometimes lower. For a period of 1 year, each hospital recorded the day on which more than 60 percent of the babies born were boys. Which hospital do you think recorded more such days?

- The larger hospital (21)
- The smaller hospital (21)
- About the same (that is, within 5 percent of each other) (53)“

Die Zahl in Klammern hinter den Antwortmöglichkeiten zeigt die Zahl der Studenten, die sich für diese Antwort entschieden haben. Die meisten schätzten die Wahrscheinlichkeit, dass es zu mehr als sechzig Prozent neugeborenen Jungen kommt, bei beiden Krankenhäusern gleich hoch ein. Kahneman und Tversky (Tversky et al. 1974, 1125) nehmen an, dass diese Meinung daher rührt, dass

⁴ Aus der ihnen eigenen, schier unermesslich scheinenden Fülle an Experimenten haben Kahneman und Tversky noch weitere Beispiele für das Auftreten der verschiedenen Verzerrungen vorzuweisen. Uns soll hier jeweils eines genügen. Für weitere Beispiele siehe beispielsweise Tversky und Kahneman 1974.

diese Ereignisse in derselben Statistik geführt werden und daher gleich repräsentativ für die Gesamtbevölkerung sind. Das heißt, die Versuchsteilnehmer halten auch in einer kleinen Stichprobe das für wahrscheinlich, was repräsentativ für die gesamte Bevölkerung ist. Tatsächlich jedoch ist die erwartete Anzahl an Tagen mit mehr als sechzig Prozent neugeborenen Jungen für das kleine Krankenhaus höher als für das große, da es für große Probengrößen weniger wahrscheinlich ist, von den fünfzig Prozent abzuweichen. Diese fundamentale statistische Auffassung ist offenbar nicht Teil der Intuitionen der Leute, meinen Kahneman und Tversky.

Eine damit verwandte Verzerrung findet sich in einer falschen Vorstellung vom Zufall (*Misconceptions of Chance*): Menschen erwarten, dass sich in einer durch einen Zufallsprozess generierten Sequenz Charakteristiken dieses Prozesses wiederfinden, auch wenn die Sequenz nur kurz ist. Betrachtet man beispielsweise eine Sequenz von Münzwürfen, so schätzen Versuchspersonen die Sequenz K/Z/K/Z/Z/K⁵ als wahrscheinlicher ein als die Sequenz K/K/K/Z/Z/Z, die nicht so zufällig erscheint, und auch als wahrscheinlicher als die Sequenz K/K/K/K/Z/K, die nicht die Gleichförmigkeit der Münze repräsentiert. (Tversky et al. 1974, 1125). Ein weiteres, recht anschauliches Beispiel für diese Verzerrung ist die *Gambler's Fallacy*. Nachdem an einem Roulettetisch eine ganze Reihe roter Zahlen hintereinander im Rouletterad ausgelost wurde, glauben die meisten Leute fälschlicherweise, es müsse nun eine schwarze Zahl an der Reihe sein. Kahneman und Tversky führen das auf den Glauben der Menschen zurück, dass das Auftreten einer schwarzen Zahl eine repräsentativere Sequenz produzieren würde als das erneute Auftreten einer roten Zahl. Der Zufall wird als ein sich selbst korrigierender Prozess angesehen, bei dem Ausschläge in das eine Extrem mit Ausschlägen in das andere Extrem ausgeglichen werden, um das Gleichgewicht wiederherzustellen. Tatsächlich jedoch werden Abweichungen nicht korrigiert, sondern mit längerer Prozesslaufzeit abgeschwächt. (Tversky et al. 1974, 1125). Diese Verzerrung konnten Kahneman und Tversky nicht nur bei Laien, sondern auch bei mit dem Phänomen vertrauten Psychologen beobachten, deren Antworten in Versuchen die Erwartung reflektierten, dass eine gültige Hypothese über eine Bevölkerung durch ein statistisch signifikantes Ergebnis in einer Stichprobe repräsentiert sein würde, ungeachtet wie groß die Stichprobe wäre. Dadurch setzen die Forscher zu hohe Erwartungen in Untersuchungsergebnisse kleiner Gruppen und überbewerten die Reproduzierbarkeit der Ergebnisse. Kahneman und Tversky (Tversky et al. 1974, 1126) sehen darin die Gefahr, dass

⁵ K = Kopf; Z = Zahl

in der Forschung Stichproben von ungeeigneter Größe ausgewählt und Resultate überinterpretiert werden.

Eine weitere Verzerrung zeigt auf, dass Menschen unter Nutzung der Repräsentativheuristik nicht nur fehlerhafte Vorhersagen treffen, sondern auch dort Vorhersagen treffen, wo gar keine Vorhersage möglich ist (*Insensitivity to Predictability*). In einer Studie bekamen die Versuchspersonen kurze Texte zu lesen, die die Leistung von Lehramtsstudenten bei der Durchführung einer Übungsschulstunde beschrieben. Ein Teil der Versuchspersonen wurde gebeten, die Qualität der beschriebenen Schulstunde in Prozentpunkten zu bewerten. Der andere Teil der Versuchspersonen wurde gebeten, die Leistung jedes Lehramtsstudenten in fünf Jahren vorherzusagen. Die Einschätzungen beider Versuchsgruppen waren identisch, das heißt, die Vorhersage eines entfernten Ereignisses (Leistung in fünf Jahren) war identisch mit der Bewertung der Information, auf der die Vorhersage beruhte (Leistung in der Übungsstunde). Kahneman und Tversky (Tversky et al. 1974, 1126) sind sich sicher, dass sich die Versuchspersonen bewusst waren, dass die Leistung eines Lehrers kaum anhand der Leistung während einer Übungsstunde fünf Jahre zuvor vorhersagbar ist. Die Versuchspersonen halten jedoch an ihren Vorhersagen fest und haben die falsche Vorstellung, dass die Vorhersagen, die sie mithilfe der Repräsentativheuristik treffen, stichhaltig sind (*Illusion of Validity*). Diese Verzerrung zeigt sich in dem oben besprochenen Beispiel von Nachbar Steve, von dem eine Beschreibung vorliegt und dessen Beruf eingeschätzt werden soll. Bei diesem Versuch steigt das Vertrauen in die eigene Vorhersage mit dem Grad der Repräsentativität, das heißt, die Leute sind stark davon überzeugt, dass eine Person ein Bibliothekar ist, wenn die vorliegende Beschreibung seiner Persönlichkeit stark dem Stereotyp eines Bibliothekars entspricht, selbst wenn die Beschreibung spärlich und unzuverlässig ist. Diese Verzerrung sehen Kahneman und Tversky (Tversky et al. 1974, 1126) auch bei Psychologen, die sich dieses Effekts bewusst sind, sich aber dennoch bei Probandenauswahlgesprächen stark von ihren Vorhersagen über die Kandidaten überzeugt zeigen, obwohl sie um die hohe Fehlbarkeit von Vorhersagen aufgrund solcher Gespräche wissen.

In der letzten Verzerrung, die wir im Zusammenhang mit der Repräsentativheuristik betrachten, geht es um falsche Vorstellungen über Regression mit der bekannten Fehleinschätzung der Regression zum Mittelwert (*Regression to the Mean*). In seinem Buch *Thinking, Fast and Slow* (Kahneman 2011, 175–176) beschreibt Kahneman die mittlerweile berühmte Geschichte einer Begebenheit, als er bei der israelischen Luftwaffe ein Seminar über Psychologie in Trainingsabhielt. Kahneman unterrichtete dort, dass die Belohnung von Leistungssteigerungen effektiver wirke als die Bestrafung von Fehlern. Ein Ausbilder der Luftwaffe

entgegnete, dass seiner Erfahrung nach Kadetten, die nach einer guten Leistung gelobt wurden, bei der nächsten Durchführung des Manövers eine schlechtere Leistung ablieferten. Wenn er dagegen diejenigen, die eine schlechte Leistung erbracht hatten, lautstark dafür kritisiert hatte, zeigten diese bei der folgenden Durchführung eine bessere Leistung. Daraus folgerte der Ausbilder, dass Bestrafung einen besseren Effekt auf die Leistung der Kadetten habe als Lob. Kahneman bemerkt, dass der Ausbilder recht hatte, zugleich aber auch unrecht. Recht hatte er mit seiner Beobachtung, dass auf eine gute Leistung eines Kadetten mit hoher Wahrscheinlichkeit beim nächsten Mal eine schlechtere Leistung folgte, während auf eine schlechte Leistung meist eine bessere folgte. Unrecht hatte der Ausbilder aber darin, dass er einen kausalen Zusammenhang zwischen der Leistungsverschlechterung beziehungsweise verbesserung der Kadetten und seinem eigenen Lob- und Tadelverhalten angenommen hatte. Für Kahneman ist klar, dass der Ausbilder hier eine Regression zum Mittelwert beobachtet, aber falsch interpretiert hatte. Der Ausbilder lobte seine Kadetten nur dann, wenn sie eine überdurchschnittlich gute Leistung erbracht hatten, was vermutlich nicht nur mit Talent, sondern auch mit etwas Glück zu tun gehabt hatte. Dass der Kadett dieses Glück beim nächsten Mal wieder haben würde, ist eher unwahrscheinlich, weshalb es wahrscheinlicher ist, dass er beim nächsten Versuch eine schwächere Leistung erbringen würde. Das hätte dann aber nichts mit dem Lob des Ausbilders zu tun. Ebenso schrie der Ausbilder auch nur dann einen Kadetten an, wenn dieser eine unterdurchschnittlich schlechte Leistung abgeliefert hatte, was wiederum vor allem mit dem Zufall zu tun gehabt hatte, weshalb der Kadett wahrscheinlich beim nächsten Versuch eine bessere Leistung erbringen würde, egal ob der Ausbilder ihn tadelte oder nicht. Kahneman sieht in der Leistung der Kadetten einen zufallsabhängigen Prozess, der zwangsläufig Schwankungen unterliegt. Kahneman und Tversky (Tversky et al. 1974, 1127) sind sich einig, dass Menschen diese Zufallsabhängigkeit des Prozesses und die Zwangsläufigkeit der daraus resultierenden Schwankungen oft nicht erkennen und daher andere kausale Gründe interpretieren, die aber nichts mit den beobachteten Schwankungen zu tun haben.

Kommen wir zur zweiten der drei Heuristiken, der Verfügbarkeitsheuristik. Bei der Verwendung dieser Heuristik schätzen Menschen beispielsweise die Wahrscheinlichkeit eines Ereignisses danach ein, wie leicht sie sich dessen Eintreten vorstellen können oder auch wie leicht sie sich an frühere solche Ereignisse erinnern können. Beispielsweise könnte man das Risiko für Herzanfälle bei Menschen mittleren Alters danach abschätzen, wie viele Personen mittleren Alters aus dem eigenen Umfeld schon einen Herzanfall hatten.

„Availability is a useful clue for assessing frequency or probability, because instances of large classes are usually recalled better and faster than instances of less frequent classes. However, availability is affected by factors other than frequency and probability.“⁶ (Tversky et al. 1974, 1127)

Verlässt man sich also auf die Verfügbarkeit, kommt es auch unter dieser Heuristik zu vorhersehbaren Verzerrungen, von denen wir hier drei betrachten werden.

Verzerrungen aufgrund der Abrufbarkeit von Beispielfällen (*Biases due to the Retrievability of Instances*): Wird die Größe einer Klasse danach beurteilt, wie leicht abrufbar eingetretene Fälle dieser Klasse sind, dann erscheinen diejenigen Klassen, deren Fälle leichter abrufbar sind, als größer als Klassen gleicher Frequenz, deren Fälle nicht so leicht abrufbar sind. Zur Veranschaulichung beschreiben Kahneman und Tversky (Tversky et al. 1974, 1127) einen ihrer Versuche, bei dem die Versuchspersonen Listen mit Namen bekannter Persönlichkeiten, sowohl männlich als auch weiblich, vorgelesen bekamen. Danach wurden sie gebeten, einzuschätzen, ob die Liste mehr Männer- oder mehr Frauennamen enthalten hatte. Bei diesem Versuch wurden verschiedenen Versuchsgruppen verschiedene Listen vorgelesen; auf manchen Listen waren die genannten Herren berühmter als die genannten Damen, während auf anderen Listen die Damen berühmter waren als die Herren. Die Versuchspersonen schätzten bei jeder dieser Listen diejenige Klasse (in diesem Fall das Geschlecht) als größer ein, das die berühmteren Namen beinhaltet hatte. Die bekannteren Namen hatten die Versuchspersonen im Nachhinein besser wieder abrufen können als die weniger bekannten und daher angenommen, das Geschlecht mit den leichter erinnerbaren Namen sei das stärker repräsentierte gewesen.

Eine weitere Art von Verzerrungen entsteht je nach Leistungsfähigkeit eines Suchkriteriums (*Biases due to the Effectiveness of a Search Set*). Diese Verzerrung ist nicht ganz einfach zu charakterisieren. Kahnemans und Tverskys Beispiel (Tversky et al. 1974, 1127) erwähnt Versuche, bei denen die Versuchspersonen gebeten wurden, sich vorzustellen, sie wählten ganz zufällig aus einem englischen Text ein Wort aus. Sie wurden gebeten einzuschätzen, ob es wahrscheinlicher ist,

⁶ Dieses Zitat zeigt meiner Meinung nach recht anschaulich das Problem mit den Heuristiken. Wenn wir Kahneman und Tversky Glauben schenken, verwenden wir Heuristiken jeden Tag. Sie helfen uns, mit Berechnungen im Alltag umzugehen und machen dadurch Entscheidungen leichter. Doch gleichzeitig machen wir gerade mit Hilfe dieser Erleichterungen Fehler, die wir ohne sie nicht gemacht hätten, aber ohne sie ist die Entscheidung komplexer und langwieriger, was im täglichen Leben übermäßig Ressourcen wie Zeit und kognitive Kapazitäten verbraucht, sodass unser tägliches Leben nicht aussehen würde wie unser tägliches Leben.

dass dieses Wort mit dem Buchstaben r beginnt, oder dass r der dritte Buchstabe dieses Wortes ist. Die Versuchspersonen suchen in ihrem Wortschatz nach Wörtern, die mit r beginnen und Wörtern, bei denen ein r an der dritten Stelle steht. Sie beurteilen die Häufigkeit der Wörter danach, wie leicht ihnen Wörter der beiden Kategorien einfallen. Da es aber um einiges leichter ist, nach Wörtern mit einem bestimmten Anfangsbuchstaben zu suchen, als nach Wörtern mit einem bestimmten Buchstaben an dritter Stelle, fallen ihnen auch mehr Wörter ein, die mit r beginnen als Wörter mit r als drittem Buchstaben. Die Versuchspersonen schätzen also die Wahrscheinlichkeit, dass das zufällig gezogene Wort mit einem r anfängt, größer ein. Das gleiche Ergebnis brachte dieser Versuch auch mit weiteren Konsonanten, die an dritter Stelle in einem Wort häufiger vorkommen als an erster. Es ist, meiner Meinung nach, nicht ganz einfach in Worte zu fassen, was diese Verzerrung von der oben besprochenen Verzerrung aufgrund der Abrufbarkeit von Beispielen und der Verzerrung der Vorstellbarkeit (siehe unten) unterscheidet. Die Leistung der Personen hängt hier nicht nur von ihrer Gedächtnisleistung ab, aber auch. Die Leistung hängt auch nicht nur von ihrer Vorstellungskraft ab, aber auch. Es scheint hier wirklich auf die Effektivität, nicht des Gedächtnisses oder der mathematischen Fähigkeiten, anzukommen, sondern auf die Effektivität der Suche nach einem bestimmten Kriterium in Gedächtnis und Vorstellung. Je effektiver diese Suche abläuft, desto besser die Versuchsergebnisse.

Bei den Verzerrungen der Vorstellbarkeit (*Biases of Imaginability*) spielt das Gedächtnis eine eher untergeordnete Rolle. Sie kommt in Fällen zum Vorschein, in denen es nicht darum geht, die Auftrittshäufigkeit einer Klasse im Gedächtnis wiederzufinden, sondern sie anhand bekannter Regeln selbst zu generieren. Vor eine solche Aufgabe gestellt, entwickeln Menschen gewöhnlich verschiedene Fallbeispiele und beurteilen ihre Wahrscheinlichkeiten danach, wie einfach sie zu konstruieren sind. Allerdings geht eine einfache Konstruierbarkeit nicht immer mit einer großen Häufigkeit einher, weshalb es zu der Verzerrung der Vorstellbarkeit kommen kann. Kahneman und Tversky (Tversky et al. 1974, 1127) beschreiben ein Beispielexperiment: Eine Gruppe von zehn Personen soll in verschiedene Untergruppen aufgeteilt werden, wobei die Gruppenstärke k zwischen zwei und acht Mitgliedern liegen soll. Wie viele verschiedene Untergruppen können gebildet werden? Die richtige Antwort liegt laut den Autoren beim Binomialkoeffizienten $\binom{10}{k}$, der bei $k = 5$ sein Maximum erreicht. Um ohne mathematische Berechnungen zu einer Lösung zu kommen, kann man versuchen, in seiner Vorstellung Untergruppen verschiedener Mitgliederstärke zu bilden und

ihre Anzahl danach beurteilen, wie leicht sie einem eingefallen sind. Untergruppen mit wenigen Mitgliedern (etwa zwei) sind dabei leichter vorzustellen als Untergruppen mit mehreren Mitgliedern (etwa acht). Teilt man die gesamte Gruppe im Kopf in verschiedene unabhängige Teile, so sieht man schnell, dass es einfacher ist, fünf Untergruppen mit je zwei Mitgliedern zu bilden, als zwei Untergruppen mit acht Mitgliedern. Daher werden die kleinen Untergruppen zahlreicher erscheinen als die großen. Bei ihrem Versuch, schreiben Kahneman und Tversky (Tversky et al. 1974, 1128), lag der Median der geschätzten Anzahl an Untergruppen mit zwei Mitgliedern bei siebzig, für Untergruppen mit acht Mitgliedern dagegen bei zwanzig. Die korrekte Antwort wäre in beiden Fällen 45 gewesen. Eine solche Aufteilungssituation mag reichlich abstrakt erscheinen. Doch da uns die Lebenserfahrung zeigt, dass wahrscheinlichere Ereignisse auch leichter vorstellbar sind als unwahrscheinlichere, nutzen wir die Verfügbarkeitsheuristik im Alltag. Dass die Verzerrung der Vorstellbarkeit dabei durchaus ein Problem darstellen kann, veranschaulicht ein Beispiel von Kahneman und Tversky (Tversky et al. 1974, 1128):

„Imaginability plays an important role in the evaluation of probabilities in real-life situations. The risk involved in an adventurous expedition, for example, is evaluated by imagining contingencies with which the expedition is not equipped to cope. If many such difficulties are vividly portrayed, the expedition can be made to appear exceedingly dangerous, although the ease with which disasters are imagined need not reflect their actual likelihood. Conversely, the risk involved in an undertaking may be grossly underestimated if some possible dangers are either difficult to conceive of, or simply do not come to mind.“

Nicht jeder plant in seinem Leben eine Expedition in das Amazonasgebiet, aber schon einen Urlaub zu planen, kann uns in Situationen falsch beurteilter Wahrscheinlichkeiten und Risiken bringen.

Die dritte und letzte Heuristik, die wir hier betrachten, ist die Ankerheuristik. Sie wird in Fällen verwendet, in denen Menschen zu einem Schätzwert gelangen, indem sie bei einem Anfangswert beginnen und diesen erweitern, um zum Endergebnis zu gelangen. Diese Anpassungen sind meist unzureichend, da unterschiedliche Anfangswerte zu unterschiedlichen Schätzwerten führen, die in Richtung des Anfangswertes abweichen. Diesen Effekt nennen Kahneman und Tversky (Tversky et al. 1974, 1128) Verankerung (*Anchoring*). Wie dabei schon anklingt, gibt es auch bei dieser Heuristik interessante Verzerrungen.

Unzureichende Anpassung (*Insufficient Adjustment*): Diese Verzerrung kann auftreten, wenn bei einer Schätzung von einem Anfangswert ausgegangen wird, wobei dieser Anfangswert auch Einfluss auf die Schätzung nehmen kann, wenn er

zufällig entstand und nichts mit dem eigentlichen Schätzwert zu tun hat. In einem Beispielexperiment von Kahneman und Tversky (Tversky et al. 1974, 1128) wurden die Versuchspersonen gebeten, verschiedene Schätzungen abzugeben, beispielsweise für die Anzahl afrikanischer Staaten in den Vereinten Nationen (UN). Für jede Schätzung wurde zu Beginn, unter den Augen der Versuchsteilnehmer, mit einem Glücksrad eine Zahl zwischen null und einhundert ermittelt. Die Personen sollten nun in einem ersten Schritt angeben, ob sie glaubten, die Zahl afrikanischer Staaten in der UN läge unter oder über diesem Zufallswert. In einem zweiten Schritt sollten sie abweichend von dieser Zahl nach unten oder oben ihren Schätzwert abgeben. Bei der Gruppe, die die Zahl 10 als Zufallswert vom Glücksrad zugeteilt bekam, lag der Median der Schätzwerte bei 25; bei der Gruppe, die die Zufallszahl 65 bekommen hatte, lag der Median bei 45. Die Anfangswerte, die, was die Versuchsteilnehmer wussten, lediglich Zufallszahlen waren und nichts mit der tatsächlichen Anzahl afrikanischer Staaten in der UN zu tun hatte, nahmen Einfluss auf die Schätzung, sodass unterschiedliche Anfangswerte zu unterschiedlichen Schätzwerten führten.

Eine weitere Verzerrung führt dazu, dass Menschen die Wahrscheinlichkeiten konjunktiver und disjunktiver Ereignisse falsch einschätzen (*Biases in the Evaluation of Conjunctive and Disjunctive Events*). Psychologin Maya Bar-Hillel (Bar-Hillel 1973) führte eine Studie durch, um diesen Effekt aufzuzeigen. In dieser Studie wurden den Versuchsteilnehmern Lotterien über verschiedene Ereignisse zur Wahl gestellt: (a) Ein einfaches Ereignis, wie das Ziehen einer roten Kugel aus einer Urne mit fünfzig Prozent roten und fünfzig Prozent weißen Kugeln; (b) ein konjunktives Ereignis, wie das Ziehen einer roten Kugel sieben Mal hintereinander mit Zurücklegen aus einer Urne mit neunzig Prozent roten und zehn Prozent weißen Kugeln; (c) ein disjunktives Ereignis, wie das Ziehen wenigstens einer roten Kugel in sieben Zügen, mit Zurücklegen, aus einer Urne mit zehn Prozent roten und neunzig Prozent weißen Kugeln. Die Mehrheit der Versuchsteilnehmer entschied sich für Lotterie (b) (Wahrscheinlichkeit 0.48) statt für Lotterie (a) (Wahrscheinlichkeit 0.50). Bei der Wahl zwischen den Lotterien (a) und (c) (Wahrscheinlichkeit 0.52) entschied sich die Mehrheit für Lotterie (a). In beiden Fällen entschieden sich die Teilnehmer für die Lotterie mit der geringeren Gewinnwahrscheinlichkeit. Diese und weitere Ergebnisse anderer Studien legen den Schluss nahe, dass Menschen dazu neigen, die Wahrscheinlichkeit konjunktiver Ereignisse zu überschätzen, während sie die Wahrscheinlichkeit disjunktiver Ereignisse unterschätzen. Diese Verzerrungen erklären Kahneman und Tversky anhand des Verankerungs-Effekts (Tversky et al. 1974, 1129):

„The stated probability of the elementary event (success at any one stage) provides a natural starting point for the estimation of the probabilities of both conjunctive and disjunctive events. Since adjustment from the starting point is typically insufficient, the final estimates remain too close to the probabilities of the elementary events in both cases. [...] As a consequence of anchoring, the overall probability will be overestimated in conjunctive problems and underestimated in disjunctive problems.“

Sie mahnen, dass diese Verzerrung im Alltag zu einem Problem werden kann, beispielsweise wenn ein Unternehmen die Einführung eines neuen Produktes oder die Durchführung eines Projektes plant. Solcherlei Vorhaben sind gewöhnlich aus verschiedenen Zielen zusammengesetzt und haben daher konjunktiven Charakter. Damit das Gesamtprojekt erfolgreich sein kann, muss jedes der Unterprojekte erfolgreich abgeschlossen werden. Die Neigung, die Wahrscheinlichkeit konjunktiver Ereignisse zu überschätzen, führt nach Kahneman und Tversky (Tversky et al. 1974, 1129) zu unangebrachtem Optimismus bei der Einschätzung, ob ein Projekt ein Erfolg werden wird oder nicht.

Für Kahneman und Tversky (Tversky et al. 1974, 1130) ist es nicht verwunderlich, dass Heuristiken wie Repräsentativ- oder die Verfügbarkeitsheuristik beibehalten werden, obwohl sie gelegentlich zu Fehlern in Vorhersagen oder Schätzungen führen können, denn die Heuristiken sind nützlich. Sie sind sicher, dass Menschen statistische Prinzipien nicht im Alltag lernen. Zwar begegnen sie statistischen Regeln wie der Regression zum Mittelwert auch in ihrem täglichen Leben. Doch kaum jemand schließt aus diesen Erfahrungen von sich aus auf statistische Grundsätze, weil die relevanten Erfahrungen nicht entsprechend verarbeitet werden.

„For example, people do not discover that successive lines in a text differ more in average word length than do successive pages, because they simply do not attend to the average word length of individual lines or pages. Thus, people do not learn the relation between sample size and sampling variability, although the data for such learning are abundant.“ (Tversky et al. 1974, 1130)

Damit machen sie klar, dass man realen Menschen nicht die gleichen kognitiven Fähigkeiten zuschreiben kann wie dem Menschenbild, das die Ökonomen als Grundlage für ihre Modelle nehmen. Vielmehr sollte die moderne Entscheidungstheorie ihrer Meinung nach die Forschungen über kognitive Verzerrungen berücksichtigen, da sie Auswirkungen auf die theoretische und angewandte Rolle von Wahrscheinlichkeitsbeurteilungen haben. Dabei beziehen sie sich vor allem auf die Idee der subjektiven Wahrscheinlichkeiten, die als quantifizierte

Beurteilung einer idealisierten Person betrachtet wird. Die subjektive Wahrscheinlichkeit eines bestimmten Ereignisses wird über die Art und Anzahl an Wetten über dieses Ereignis bestimmt, die eine Person bereit ist, einzugehen. Subjektiv ist diese Wahrscheinlichkeit deshalb, weil sie von Individuum zu Individuum unterschiedlich ausfallen kann. An dieser Konzeption kritisieren Kahneman und Tversky (Tversky et al. 1974, 1130), dass, obwohl subjektive Wahrscheinlichkeiten manchmal aus Präferenzen gegenüber Wetten abgeleitet werden können, sie nicht üblicherweise auch auf diesem Weg entstehen. Wenn eine Person auf Team A statt auf Team B wettet, dann tut sie das, weil sie glaubt, dass Team A die größeren Gewinnchancen hat. Sie leitet diesen Glauben aber nicht aus ihrem Wettverhalten ab. Damit wollen Kahneman und Tversky klarstellen, dass in der Realität subjektive Wahrscheinlichkeiten die Präferenzen beim Wetten bestimmen, und nicht von diesen Präferenzen abgeleitet werden, wie es die axiomatische *Rational Decision Theory* sieht. Kahneman und Tversky (Tversky et al. 1974, 1131) werben für mehr Erforschung für ein besseres Verständnis von Heuristiken und den Verzerrungen, die ihre Verwendung verursacht, denn nicht nur haben sie Auswirkungen auf die theoretische Rolle von subjektiven Wahrscheinlichkeiten, ein besseres Verständnis könnte auch zu besseren Entscheidungen unter Unsicherheit führen.

3.4 Selten und die Spieltheorie

Kommen wir zum vierten und wichtigsten Faktor in der Entstehung der Verhaltensökonomie nach Cartwright (Cartwright 2014), die Entwicklungen der Spieltheorie in der zweiten Hälfte des zwanzigsten Jahrhunderts. 1950 beschrieb der Mathematiker Nash in seiner Dissertation über nicht-kooperative Spiele das Nash-Gleichgewicht, das die Arbeit von John von Neumann und Oskar Morgenstern erweiterte und die Spieltheorie in der Ökonomie zu einem robusten und vielseitig einsetzbaren Werkzeug machte. Das Nash-Gleichgewicht wurde zu einem nützlichen Wegweiser zur rationalen Lösung nicht-kooperativer Spiele. Jedoch blieb die große Frage, wie in Situationen mit mehreren Gleichgewichten eines auszuwählen sei. Das Problem für die Ökonomen ist: Stehen mehrere Gleichgewichte zur Auswahl, so ist unklar, nach welcher Strategie der ökonomische Agent, *Homo oeconomicus*, handeln sollte. Es wird unklar, wie die rationale Handlung aussieht. Es musste also eine Methode gefunden werden, um einerseits das richtige Gleichgewicht auszuwählen, was immer „richtig“ in diesem Zusammenhang bedeuten mag. Ist ein Gleichgewicht *sinnvoller* als die übrigen oder wird es wahrscheinlicher eintreffen? Andersherum fehlte auch das Verständnis

dafür, warum Spieler in entsprechenden Situationen das eine und nicht das andere Gleichgewicht gewählt hatten. In den folgenden Jahrzehnten entwarfen zahlreiche Mathematiker und Ökonomen unterschiedliche Theorien zur Verfeinerung des Nash-Gleichgewichts. Einer davon ist Ökonom Selten. Zusammen mit Nash und Harsanyi erhielt er 1994 den Wirtschaftsnobelpreis für seine Errungenschaften auf dem Gebiet der Spieltheorie. Für die Entwicklung der Verhaltensökonomie war dabei wichtig, dass der große Beitrag, den Harsanyi und Selten zur Spieltheorie erbrachten, auf der Idee beruhte, psychologische Ideen hinzuzuziehen und ökonomische Experimente durchzuführen, um der Frage nachzugehen, wie Probanden in Spielen mit mehreren Gleichgewichten entscheiden würden.

Ende der 1950er Jahre war Selten einer der ersten Ökonomen, die Experimente durchführten, Mitte der 1980er Jahre richtete er das erste voll rechnergestützte Labor Europas ein, das sich ausschließlich mit der Erforschung ökonomischer Fragen beschäftigte (Ockenfels et al. 2010, v). Die ersten Experimente in den 1950er Jahren drehten sich vor allem um das Verhalten von Probanden in Oligopolmärkten. Offenbar (Alt et al. 1999, 253) war die damals gängige Ansicht, Oligopolisten kämen über das gegenseitige Senden von Signalen zu einer stillschweigenden Preiskoordination, die zu einem Pareto-optimalen Ergebnis führt. Werfen wir einen kurzen Blick auf eines der ersten Experimente, die er zusammen mit Heinz Sauer mann an der Universität Frankfurt a. M. durchführte. Nicht nur aus historischem Interesse ist das Experiment interessant, das sich noch stark von heutigen Versuchsaufbauten unterscheidet:

Dem Versuch liegt ein Modell mit drei Firmen zugrunde (Sauer mann et al. 1959, 428), die ein homogenes Gut anbieten. Die Kostenverläufe und Produktionskapazitäten sind für jede Firma festgelegt, unterscheiden sich jedoch untereinander. Es ist ein wenig unübersichtlich, welche Informationen den Firmen bekannt sind und welche nicht. Manche Informationen können sie einkaufen, andere nur durch Erfahrung erschließen. Jede Firma kennt die Kapazitäten der übrigen, aber nicht deren Kostenverläufe. Ihr eigener Kostenverlauf ist ihnen natürlich bekannt. Sie kennen auch nicht die Preis-Absatz-Funktion, die zeitu nabhängig jeder Gesamtangebotsmenge einen Preis zuordnet, können sie aber im Laufe des Versuchs erschließen. Informationen kaufen können die Firmen zu Beginn jeder Handelsperiode über die Schulden oder den Marktanteil der übrigen Firmen in der vorangegangenen Periode. Hierfür, oder zur Tilgung ihrer Produktionskosten, können die Firmen Kredite aufnehmen, welche Zinskosten verursachen.

Zum Ablauf: Zu Beginn einer Handelsperiode können die Firmen Informationen über die Konkurrenten einkaufen, wenn sie sich dazu entschließen. Danach

bestimmt jede Firma ihre Produktionsmenge für die aktuelle Periode. Die Produktionskosten werden sofort fällig. Diese können aus dem eigenen Barbestand oder aus Neuverschuldung finanziert werden. Eine Lagerhaltung für spätere Perioden ist nicht vorgesehen, alles, was produziert wird, kommt in derselben Periode auf den Markt und wird verkauft. Die Preis-Absatz-Funktion ist annähernd konkav, mit Abweichungen, um eine Extrapolation auf ihren Verlauf aus den Daten der bisherigen Perioden zu erschweren. Die Probanden wissen allerdings, dass ihre Gestalt an eine Hyperbel erinnert und der Preis sinkt, je höher das Gesamtangebot ausfällt (Sauermann et al. 1959, 430). Am Ende der Periode wird den Firmen ihr jeweiliger Umsatz ausbezahlt, wodurch neues Kapital für die nächste Periode bereitsteht.

Jeder der dreizehn von Sauermann und Selten durchgeführten Versuche hatte 30 Perioden und dauerte circa vier Stunden mit einer ausführlichen Einführung der Versuchsteilnehmer. Diese kamen vor allem aus den Erst- bis Drittsemestern der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität Frankfurt. Jeder der drei Firmen wurden fünf bis sechs Personen zugeteilt, von denen zwei willkürlich ausgewählt und zum Unternehmer beziehungsweise Bankagenten bestimmt wurden. Die Unternehmer treffen die Entscheidungen für ihre Firma, nach Diskussion mit den übrigen Firmenmitgliedern. Der Bankagent hält die Verbindung mit der Versuchszentrale, die die Spielleitung innehat und ihrerseits mit mindestens zwei Leuten besetzt ist. Um Verbindungen mit den übrigen Firmen zu vermeiden, ist jede Firma in einem eigenen Raum untergebracht und kommuniziert über Telefone mit der Zentrale. Während einer Periode telefoniert jede Firma drei Mal mit der Zentrale. Dabei werden entweder von der Firma oder der Zentrale verschiedene Größen durchgegeben (Sauermann et al. 1959, 433):

1. Informationen
2. Schuldenstand, Angebot
3. Preis, Gesamtangebot, Einzelumsatz

Jede dieser Größen wird von je zwei Protokollanten in den Firmen festgehalten. Dafür erhalten sie eine vorgefertigte Tabelle, in die sie für jede Periode Daten zu Preis und Gesamtangebot, sowie die Einzelangebote jeder Firma festhalten. Neben dem eigenen Schuldenstand werden auch die gegebenenfalls eingekauften Informationen über die Schulden der übrigen Firmen aufgeschrieben. Die eigenen Produktionskosten und der Periodenumsatz werden in der Tabelle vermerkt, um daneben den Rohgewinn angeben zu können, von dem noch die zu zahlenden Zinsen und Informationskosten abgezogen werden müssen. Der entstehende Reingewinn plus dem Vermögensstand aus der vorangegangenen Periode ergeben

den Vermögensstand der Firma. Zusätzlich werden die Probanden gebeten, die Gründe und vorgebrachten Argumente für ihre Entscheidungen in den einzelnen Perioden stichwortartig zu protokollieren und dieses Protokoll nach Versuchsende einzureichen. In den Versuchen wird mit Spielgeld gehandelt, um die Kosten und Geldflüsse zu verdeutlichen.

Der Preis auf dem Versuchsmarkt wird über das Gesamtangebot mithilfe der Preis-Absatz-Funktion ermittelt. Natürlich gibt es dabei Angebotskombinationen der drei Versuchsfirmen, die einen höheren oder einen niedrigeren Gesamtumsatz erzielen. Sauermann und Selten (Sauermann et al. 1959, 445–450) errechnen die Pareto-optimalen Angebotskombinationen aus allen theoretisch möglichen Kombinationen, denn diese würden erzielt werden, wenn sich die Firmen mit stillschweigenden Übereinkünften absprächen, wie es die gängige Hypothese annimmt. Die Ergebnisse vergleichen sie mit den erzielten Daten aus den Versuchen: In insgesamt 390 Perioden kam es lediglich in dreien zu solchen Pareto-optimalen Kombinationen.

„Das zeigt, daß stillschweigende Übereinkommen auf der Basis *Pareto*-optimaler Punkte in der vorliegenden nicht kooperativen Situation kaum zu erwarten sind. Damit werden in diesem Zusammenhang auch Annahmen über gemeinsame Gewinnmaximierung der Unternehmungen höchst unrealistisch.“ (Sauermann et al. 1959, 450)

Vielmehr sehen sie (Sauermann et al. 1959, 463) eine Anpassung an das erwartete Konkurrenzangebot im Sinne von Cournot, wo die Angebotsmenge durch wechselseitige Überlegungen über die Angebotsmenge der Konkurrenzfirmen und die eigene beste Antwort auf die erwarteten Mengen zustande kommt. Auf diese Weise wird in der Cournotschen Theorie eine stabile Gesamtangebotsmenge, das Cournot-Gleichgewicht, erreicht.⁷ Es scheint also, als seien es Überlegungen, wie Selten sie aus der Spieltheorie kannte, die die Probanden zu ihren Entscheidungen brachten. Diese Beobachtung wiederholte Selten in weiteren Oligopolexperimenten. Bei einem Experiment mit Nachfrageträgheit war nicht ganz klar, wie die Entsprechung zur Cournot-Lösung aussehen würde. Auf der Suche nach einer Lösung dieses dynamischen Spiels fand er eine Gleichgewichtslösung mit Hilfe der Rückwärtsinduktion. Doch das Spiel hatte noch weitere Gleichgewichtspunkte, wodurch die Frage aufkam, was die eine Lösung von

⁷ Das ist eine sehr grobe Beschreibung von Cournots Theorien zum Oligopolmarkt. Sie soll an dieser Stelle aber für unsere Zwecke genügen. Für ein genaueres Studium sei empfohlen: Cournot, A. A. (1838) *Recherches sur les principes mathématiques de la théorie des richesses*, Paris: Imprimerie d'Amédée Gratiot, bes. S. 88–100.

den übrigen Gleichgewichtspunkten unterschied. An dieser Stelle führte er die Subspiel-Perfektion als generelles Konzept der Spieltheorie ein, neben der Arbeit mit Harsanyi ein weiterer Zugewinn für die Spieltheorie, der heute aus keinem Lehrbuch mehr wegzudenken ist. (Albers et al. 1997, 5).

Kurz nachdem Selten mit seinen Experimenten zur Oligopoltheorie begonnen hatte, begann er auch, Werke von Herbert Simon zur Idee der *bounded Rationality* zu lesen, die ihn von der Notwendigkeit einer Theorie der eingeschränkten Rationalität überzeugten:

„Shortly after I had began (sic!) to run oligopoly experiments I read Herbert Simon and I was immediately convinced of the necessity to build a theory of bounded rationality. At first I still thought it might be possible to do this in an axiomatic way, just by pure thinking, as you build up a theory of full rationality. But over the years I more and more came to the conclusion that the axiomatic approach is premature. Behavior cannot be invented in the armchair, it must be observed in experiments. You first have to know what really happens before forming an axiomatic theory.“ (Albers et al. 1997, 4)

Es ist deutlich, welche Bedeutung der Mathematiker Selten Laborexperimenten beimisst. Mit dieser Einstellung beförderte er die Bedeutung und Entwicklung von Laborexperimenten in der ökonomischen Forschung in Europa, sowie Kollegen wie Smith es in den USA taten.

Ähnlich wie bei Selten erfuhren die Ansichten von Wissenschaftlern wie Simon und Kahneman und Tversky in den vergangenen Jahrzehnten auch unter Ökonomen mehr und mehr an Zustimmung, die nicht zuletzt durch die steigende Popularität von Laborexperimenten gefördert wurde. Zwischenzeitlich konnten Ökonomen an ihrer eigenen Arbeit sehen, dass Theorien mit naiven Rationalitätsannahmen nicht das in den Laboren beobachtete Verhalten von Probanden widerspiegeln, geschweige denn korrekt vorhersagen konnten. Das machte deutlich, dass alternative Theorien gefunden werden mussten. Die Bereitschaft, über den bisherigen Tellerrand hinwegzuschauen, war nun vorhanden, doch was nutzte das, ohne die passenden Werkzeuge? Für die Entstehung der modernen Verhaltensökonomie genühten diese ersten drei der besprochenen Aspekte noch nicht. Der Vierte brachte die Lösung: Er brachte die Spieltheorie als wichtiges und erfolgreiches Werkzeug für die Entwicklung und Analyse der Laborexperimente. Ohne ein solches Werkzeug sind die Experimente nicht von großem Nutzen für die Wissenschaftler.⁸ Vor der Entwicklung, die sie durch die Arbeit vor allem

⁸ Diese Auffassung ist ja auch einer der Gründe, weshalb Neurowissenschaftler sich für Neuroökonomie interessieren.

von Nash, Selten und Harsanyi erfahren hat, wäre sie kein so leistungsstarkes Werkzeug gewesen, um die Verhaltensökonomie voranzutreiben. Interessant dabei ist, dass vor allem Selten seinen Beitrag zur Spieltheorie vor allem und erst im Zusammenhang mit Laborexperimenten erbracht hat.

Wir haben also in den 1950er Jahren die Auffassung, dass eine absolute Rationalität, wie sie die Ökonomie von ihren Agenten fordert, von gewöhnlichen Menschen, also den ökonomischen Agenten der realen Welt, nicht erfüllt wird. Dies wird vor allem von Forschern außerhalb der Ökonomie so gesehen. Ungefähr zur gleichen Zeit verbreiten sich Laborexperimente als interessantes neues Betätigungsfeld unter Ökonomen und in ihrem Anhang auch die Ahnung, dass die Warnungen der Sozialwissenschaftler und Psychologen vor einer ökonomischen Theorie mit einem allzu rationalen homo oeconomicus doch nicht so ohne Weiteres beiseite gewischt werden könnten. In den 1960er bis 1980er Jahren machen die Entwicklungen in der Spieltheorie Laborexperimente zu einem ernstzunehmenden Verfahren in der ökonomischen Forschung. In den Augen von Verhaltensökonomern sind sie ein geeignetes Hilfsmittel, um Entscheidungsfindungsprozesse nachvollziehen zu können. Bald werden nicht nur bestehende Theorien auf ihre Realitätsnähe getestet, sondern auch neue Einflüsse in die ökonomische Theorie. Wie schon an den Experimenten von Smith und Selten zu sehen ist, können ganz unterschiedliche Dinge im Fokus ökonomischer Experimente stehen, von ganzen Institutionen bis hin zum Verhalten einzelner Probanden.

3.5 Soziale Präferenzen

Durch unseren Einblick in die Historie der Verhaltensökonomie lässt sich nachvollziehen, wie die Verhaltensökonomie wurde, was sie heute ist. Inzwischen ist die Verhaltensökonomie ein etablierter Teilbereich der Ökonomie, der Laborexperimente als reguläres Forschungsinstrument zum Erkenntnisgewinn einsetzt.

Das generelle Ziel von Verhaltensökonomern ist es heute, menschliches Verhalten und seine Konsequenzen zu verstehen: Warum kauft jemand eine Pizza oder geht regelmäßig zur Arbeit? Warum kaufen Menschen Versicherungen, um sich gegen die geringsten Risiken abzusichern und gehen danach zum Kiosk, um Lotto zu spielen? Warum lassen sie sich von veganer Ernährung überzeugen, können aber nicht mit dem Rauchen aufhören? Sind das gute oder schlechte Handlungen und wie könnte man ihnen dabei helfen, bessere Entscheidungen zu treffen? Eigentlich besitzt die Ökonomie ihr berühmtes Standardmodell für solche

Fragen, doch wie wir auf den vergangenen Seiten sahen, sind sich die Verhaltensökonominnen bewusst, dass das ökonomische Standardmodell seine Schwächen hat. Allerdings hat es auch seine Stärken, derer sie sich ebenso bewusst sind. Es kann gute Vorhersagen darüber machen, welche Wahl ein Proband treffen wird. Sein axiomatischer Zugang ohne spekulative Nebenannahmen hat großen Reiz für Ökonomen. Daher verwerfen sie das ökonomische Standardmodell nicht, sondern versuchen, es durch verschiedene Änderungen besser an beobachtetes Verhalten anzupassen. Denn die Verhaltensökonominnen sind sich inzwischen sicher, dass reale Menschen nicht rational denken, wie es das ökonomische Modell ihnen zuschreibt. Vielmehr sind sie überzeugt, dass reale Menschen sich auf Heuristiken verlassen, wie sie von Kahnemann und Tversky formuliert wurden und durchaus auch auf die kognitiven Verzerrungen hereinfließen. Diese Fallibilität muss in den Veränderungen am Standardmodell berücksichtigt werden. Zu den Änderungen gehört, die Annahmen über Präferenzen ökonomischer Agenten auszubauen, um sie in deren Nutzenfunktionen einbauen zu können. Der Pool möglicher Präferenzen wird vergrößert, um das Verhalten realer Menschen zutreffender zu beschreiben. Dazu zählen Zeitpräferenzen, oder Annahmen über das Verhalten von Individuen bei Entscheidungen unter Unsicherheit, wie es Kahnemann und Tversky untersucht haben.

Ein wichtiges Beispiel sind soziale Präferenzen. Sie werden Individuen zugeschrieben, um beobachtetes Verhalten modellieren zu können. Sie sind eine Antwort der Verhaltensökonomie auf beobachtete Abweichungen vom Standardmodell. In diesem sind ökonomische Agenten ausschließlich von Egoismus motiviert. Verhaltensökonominnen zeigen sich damit nicht einverstanden angesichts einer ganzen Reihe an Beispielen für Reziprozität im Verhalten von ökonomischen Agenten. Abgesehen davon, dass eine vollständige Gleichgültigkeit gegenüber den Konsequenzen der eigenen Handlung für andere Menschen tatsächlich nur für pathologische Agenten möglich ist, gibt es empirische Evidenz gegen das Modell rein egoistisch motivierter Handlungen (Wilkinson 2008, 327) wie beispielsweise Spenden für wohltätige Zwecke, freiwillige unbezahlte Arbeit, Beteiligung an der Bereitstellung öffentlicher Güter, Bestrafung von Trittbrettfahrern, auch wenn dies mit eigenen Kosten verbunden ist, Vertrauen in Vertrauensspielen, großzügige Angebote in Ultimatumspielen, Ablehnen niedriger Angebote in Ultimatumspielen.

Neben alltäglichen Situationen sind das Evidenzen aus Experimenten. In Abschnitt 2.2 haben wir das Ultimatumspiel, das Vertrauensspiel und öffentliche Güter bereits kennengelernt. Das Verhalten von Probanden in diesen Spielen weicht regelmäßig zum Teil stark von den Vorhersagen des Standardmodells ab. Rufen wir uns das noch einmal kurz ins Gedächtnis, ohne Anspruch

auf Vollständigkeit oder die Betrachtung möglicher Motive und Auswirkungen von Variationen im Experimentaufbau: Im Ultimatumspiel zeigt, wie Wilkinson (2008, 333) zusammenfasst, eine Reihe an Studien, dass zwischen 60% und 80% der Angebote zwischen 40% und 50% der zur Verfügung stehenden Geldsumme liegen, wogegen Angebote unter 20% selten vorkommen. Auf der anderen Seite werden Angebote unter 20% in ungefähr der Hälfte der Fälle abgelehnt. Dies wird als Evidenz dafür interpretiert, dass Personen in der Rolle von Spieler 2 von Angeboten verärgert sind, die sie als unfair betrachten, und dazu bereit sind, solch unfaires Verhalten zu bestrafen, auf wenn dadurch Kosten für sie selbst entstehen. In Vertrauensspielen zeigen Probanden Reziprozität, was vom ökonomischen Standardmodell nicht vorhergesagt wird. In einer Studie von Berg, Dickhaut und McCabe (1995) zeigt sich, dass lediglich 6% der Investoren nichts, also einen Betrag von 0 investieren. Dagegen investiert der Durchschnitt der Investoren ungefähr 50% ihres Gesamtbetrags, 16% gar den gesamten Betrag. Im Durchschnitt wurden circa 95% des investierten Betrags vom Treuhänder zurückgegeben, wobei es hier starke Variationen gibt. Auch in Spielen um öffentliche Güter gibt es stark variierende Antworten, manche Spieler geben nichts, andere ihre gesamten Ressourcen für die Bereitstellung eines öffentlichen Guts. Generell, fasst Wilkinson (2008, 338) zusammen, stellen Probanden jedoch circa die Hälfte ihres Betrags zur Verfügung.

Wie gesagt, ist eine Idee der Verhaltensökonomie, die Anomalien in Standardmodellen durch Erweiterungen des Modells abzudecken. Um die evidente nicht ausschließlich egoistische Motivation von Individuen zu erklären, wurden soziale Präferenzen modelliert. Zu ihnen gibt es mehrere Modelle mit unterschiedlichen Ansätzen. Wir werden kurz zwei der bekanntesten Modelle betrachten, um die Ideen dahinter zu verstehen: Das Fehr-Schmidt-Modell und das Rabin-Modell.

Das Fehr-Schmidt-Modell ist ein Ungleichheitsaversionsmodell. Diese zeichnen sich dadurch aus, dass sie annehmen, Individuen interessieren sich für ihren eigenen Payoff, aber auch für den Payoff einer bestimmten Referenzgruppe, beziehungsweise wie sich die Höhe des eigenen Payoffs gegenüber der Payoffs anderer Individuen verhält. Die Frage, ob andere Personen mehr oder weniger Payoff erhalten als sie selbst, spielt in diesen Modellen eine Rolle für das Verhalten von Individuen, im Gegensatz zum Modell von rein egoistisch motivierten Agenten. In Ungleichheitsaversionsmodellen interessieren sich die Agenten dafür, wie ungleich die Payoffs sind; sie vergleichen ihren eigenen mit denen, die mehr oder weniger erhalten als sie selbst. Um das beschreiben zu können, werden Präferenzen eingeführt, die modellieren, inwiefern sich Individuen für die Payoffs ihrer Mitspieler interessieren. Betrachtet man das Verhalten von Versuchspersonen in Ultimatum- oder Diktatorspielen, die (hohe) Angebote machen, könnte

Altruismus eine Rolle spielen. Dieser allein wird aber nicht ausreichen, um das Verhältnis zu Mitspielern zu beschreiben, wenn man das Verhalten von Probanden in anderen Spielen betrachtet, wie beispielsweise in Vertrauensspielen oder Spielen mit öffentlichen Gütern mit Bestrafungsmöglichkeiten, in denen Versuchspersonen unkooperative Mitspieler bestrafen, auch wenn das mit Kosten für sie selbst verbunden ist. Das könnte dafür sprechen, dass Gefühle wie Neid eine Rolle spielen. Tatsächlich bindet das Fehr-Schmidt-Modell Egoismus, Altruismus und Neid ein. Dazu erweitert es die Nutzenfunktion von Spielern um einen Koeffizienten α_i für die Abneigung des Spielers i gegenüber nachteiliger Ungleichheit, sowie einen Koeffizienten β_i für seine Abneigung gegenüber vorteilhafter Ungleichheit. Sie werden manchmal auch als Koeffizienten für Neid und Schuld bezeichnet. Ihre Höhe entscheidet darüber, wie sich ein Spieler in bestimmten Situationen gegenüber den Mitspielern verhält. Nehmen wir als Beispiel ein Spiel mit öffentlichen Gütern mit Bestrafung, bei dem Spieler i sich finanziell beteiligt, und erkennt, dass Mitspieler das nicht tun. Ist α_i groß genug, wird Spieler i die nicht-kooperativen Mitspieler bestrafen. Ist andererseits β_i groß genug, wird der Spieler zum öffentlichen Gut beitragen. Studien haben gezeigt, dass solche Bestrafungen zu mehr Kooperation in Spielen mit öffentlichen Gütern führen. (Wilkinson 2008, 354–357; Dhami 2016, 399–409). Damit kann das Fehr-Schmidt-Modell das Verhalten von Probanden im normalen Ultimatum- und Diktatorspiel sowie in Spielen mit öffentlichen Gütern erklären. (Wilkinson 2008, 355). Das Fehr-Schmidt-Modell hat aber auch eine charakteristische Besonderheit, die es von anderen Modellen sozialer Präferenzen unterscheidet: Im Fehr-Schmidt-Modell vergleichen die Individuen zwar ihre eigenen Payoffs mit denen ihrer Mitspieler. Im Gegensatz zu Reziprozitätsmodellen wie dem Rabin-Modell modelliert das Fehr-Schmidt-Modell aber nicht die Einschätzungen von Spielern über die Intentionen ihrer Mitspieler. Diese spielen hier keine gesonderte Rolle. Das Fehr-Schmidt-Modell hat allerdings den Vorteil, mit seiner Beschränkung auf zwei Parameter in einer linearen Nutzenfunktion einfach aufgebaut zu sein und auf eine große Bandbreite an Spielsituationen angewendet werden zu können.

Das Rabin-Modell ist, wie gesagt, ein Reziprozitätsmodell. Es geht von folgender zentraler Annahme aus, die Wilkinson (Wilkinson 2008, 359) zusammenfasst:

„If somebody is being nice to you, fairness dictates that you be nice to him. If somebody is being mean to you, fairness allows – and vindictiveness dictates – that you be mean to him.“

Faires Verhalten soll also mit Fairness beantwortet werden, während unfaires Verhalten mit einer unfairen Antwort bedacht werden darf. Das Rabin-Modell geht daher davon aus, dass für die meisten Leute Fairness nicht nur etwas mit Gleichheit beziehungsweise Ungleichheit zu tun hat, sondern auch mit der Absicht, die hinter den Handlungen von Personen steckt. Dabei spielt die Reziprozität eine große Rolle, also ob die eigenen Intentionen vom Gegenüber gespiegelt werden. Reziprozität wird häufig in empirischen Untersuchungen beobachtet, wenn beispielsweise Spieler 2 die Freundlichkeit von Spieler 1 mit Freundlichkeit erwidert (positive Reziprozität), oder wenn Spieler 2 die Unfreundlichkeit von Spieler 1 mit Unfreundlichkeit beantwortet (negative Reziprozität). Jedoch kann Altruismus, eine Herangehensweise anderer Modelle, die negative Reziprozität nicht erklären.

Das bedeutet für Rabins Zwei-Spieler-Modell, dass der Nutzen und die Strategie der Spieler von ihrer Meinung über die Absichten des anderen Spielers abhängen. Die Strategie a_1 von Spieler 1 beruht auf seinen Annahmen b_2 über die Strategie von Spieler 2 und seinen Annahmen c_1 über die Annahmen von Spieler 2 über die Strategie von Spieler 1. Das sind zwei verschiedene Konstrukte, die in die Nutzenfunktion des Rabin-Modells einfließen: (1.) die Freundlichkeit von Spieler 1 gegenüber Spieler 2 sowie (2.) die Wahrnehmung der Freundlichkeit von Spieler 2 durch Spieler 1. Die Freundlichkeit von Spieler 1 gegenüber Spieler 2 wird durch eine Funktion von der Strategie a_1 und den Annahmen über die Strategie von Spieler 2, b_2 , abgebildet. Die Wahrnehmung der Freundlichkeit von Spieler 2 durch Spieler 1 wird abgebildet durch eine Funktion von b_2 und den Annahmen von Spieler 1 über die Annahmen von Spieler 2 über die Strategie von Spieler 1, c_1 . Daraus ergibt sich eine dreiteilige Nutzenfunktion zusammen mit dem Payoff von Spieler 1, der von a_1 und b_2 abhängt. Der entstehende Term ist sowohl dann positiv, wenn die Großzügigkeit von Spieler 1 durch Spieler 2 erwidert wird, als auch wenn seine Unfreundlichkeit von Spieler 2 erwidert wird. Somit entsteht im Rabin-Modell für positive wie auch für negative Reziprozität ein positives Nutzenergebnis, im Gegensatz zu Altruismus-Modellen. (Wilkinson 2008, 360).

Dies sind nur zwei Modelle für soziale Präferenzen, es wurden noch eine Reihe weiterer Modelle vorgeschlagen, ebenso wie beispielsweise für hyperbolisches Diskontieren oder für Entscheidungen unter Risiko, um nur zwei zu nennen. Das zeigt, dass es viele Ideen gibt, wie menschliches Verhalten in der Ökonomie modelliert werden könnte. Unser historischer Abriss zeigt, dass diese Ideen nicht ohne empirische Datenbasis zustande kommen können. Laborexperimente sind daher eines der wichtigsten Werkzeuge der modernen Verhaltensökonomie, denn sie liefern Belege für menschliches Verhalten. Weitere Hilfsmittel sind Einsichten

aus der Psychologie oder anderer Sozialwissenschaften. Wie passt die Neuroökonomie, um die es uns ja geht, da hinein? Die Neuroökonomie ist in dieser Gruppe so etwas wie der jüngste Kooperationspartner. Psychologische Modelle wurden aufgegriffen, weil sie sich, im Gegensatz zum ökonomischen Standardmodell, nicht, oder nicht nur, mit der beobachtbaren Wahl eines ökonomischen Agenten beschäftigen. Sie fragen nach den Vorgängen, die die Handelnden durchlaufen, um zu einer Entscheidung zu kommen. Zudem sind Psychologen viel mehr als Ökonomen mit der Planung, Durchführung und Auswertung von Experimenten vertraut. Bereits während ihrer Ausbildung kommen sie damit vielfach in Kontakt und noch heute ist der Anteil an Psychologen, die Erfahrung mit Experimenten haben, deutlich höher als der Anteil an Ökonomen (Cartwright 2014, 11). Das macht sie für Verhaltensökonomern zu wertvollen Partnern bei der Erforschung menschlichen Verhaltens. Das trifft auch auf kognitive Neurowissenschaftler zu. Allerdings heißt das nicht, dass auf psychologischen Modellen basierende Annahmen generell besser für die ökonomische Forschung sind. Denn anders als den Psychologen ist es Ökonomen eher mittelbar wichtig, welche Prozesse die Menschen durchlaufen, um zu einer Kaufentscheidung oder einer unternehmerischen Strategieentscheidung zu gelangen. Am Ende müssen die Vorhersagen zutreffen. Und das ist auch bei psychologisch basierten Vorhersagen nicht immer der Fall.

Die Verhaltensökonomern suchen also weiter nach passenden Veränderungen am Standardmodell. Auch diese müssen vor allem mit akkuraten Vorhersagen aufwarten können. Vorhersagen, die besser sind als die des Standardmodells. Die Messlatte liegt hoch für jede neue Änderung.⁹ Daher werfen wir im folgenden Kapitel einen Blick auf die Neurowissenschaften, die neue Erkenntnisse und Ideen in die Verhaltensökonomie einbringen möchten. Wir werden sehen, welche Geschichte und Erfahrungen sie mitbringen und welche Fragen sie zu beantworten versuchen.

⁹ Dieser kurze Einblick in die Ziele und Ansichten der Verhaltensökonomie sollen uns an dieser Stelle genügen. In den folgenden beiden Kapiteln werden wir noch über Beispiele verhaltensökonomischer Experimente sprechen. Kontroversen und Strömungen innerhalb der Verhaltensökonomie spielen hier aber keine allzu große Rolle. Weitere Informationen dazu finden sich beispielsweise in Cartwright 2014 oder Heukelom 2014.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Nachdem wir in Kapitel 3 die Verhaltensökonomie betrachtet haben, werfen wir nun einen Blick auf die andere Seite der Neuroökonomie, auf die kognitiven Neurowissenschaften. Auch hier werden wir uns in Abschnitt 4.1 zunächst kurz mit der Historie der kognitiven Neurowissenschaften beschäftigen, die uns Einblick in die Entwicklung der Fragestellungen gewähren wird, denen die Neurowissenschaftler seit dem 19. Jahrhundert nachgehen. Wir werden Ideen und Methoden von Franz Josef Gall, Pierre-Paul Broca und Carl Wernicke kennenlernen und ihre Auswirkungen auf die Frage, ob das Gehirn mit lokalen Funktionen arbeitet oder die *Mass-Action*-Theorie zutrifft. Im 20. Jahrhundert ermöglichen Umbrüche den Aufstieg des Behaviorismus zur vorherrschenden Methode. Dieser Überblick wird uns dabei helfen, in Abschnitt 4.2 zu verstehen, was die kognitiven Neurowissenschaftler heute antreibt und welche Fragen sie beschäftigen. Die Überwindung des Behaviorismus ermöglichte die Entstehung der kognitiven Neurowissenschaften und ihre Erforschung von Gedächtnis, Emotionen und anderem mehr. Diese beiden modernen Forschungsgebiete werden wir genauer betrachten, da sie auch für die neuroökonomische Forschung interessant sind. Ebenso interessant für das Verständnis derzeitiger Fragen in den Neurowissenschaften sind zwei Großprojekte in Europa und den USA. Das Human Brain Project und die BRAIN Initiative beschäftigen sich mit sehr ähnlichen übergeordneten Fragen und wir werden sehen, wie sie diese beantworten möchten. Abschnitt 4.3 liefert dann einen Überblick über heute gebräuchliche empirische Methoden, mit deren Hilfe diese Fragen beantwortet werden sollen. Dazu gehören non-invasive Methoden wie funktionale Magnetresonanztomographie (fMRT), Positronenemissionstomographie (PET), Elektroenzephalographie (EEG) und die transkranielle Magnetstimulation (TMS), sowie invasive Methoden wie die lokale Ableitung von Aktionspotenzialen. Wir werden jedes dieser

Instrumente und ihre technische Funktionalität in einem kurzen Überblick kennenlernen. Diese Betrachtung der praktischen Arbeit dient als Grundlage für das Verständnis des Beitrags der kognitiven Neurowissenschaften zur Neuroökonomie. Zusätzlich fußen einige Kritiken, die gegen die Neuroökonomie vorgebracht werden, auf diesen empirischen Methoden und der Art ihrer Anwendung. Ein Verständnis über die Grundlagen der empirischen Methoden wird daher für die Erfassung und Beurteilung dieser Kritiken in Kapitel 6 hilfreich sein.

4.1 Fragen und Historie der kognitiven Neurowissenschaften

Was ist unter dem Begriff „kognitive Neurowissenschaften“ zu verstehen? Der Neuropsychologe Wolfgang Hartje (Hartje 2012, 2) erläutert, dass die kognitiven Neurowissenschaften, wie auch andere Disziplinen der Neurowissenschaften, die Aufklärung der Strukturen und Funktionsweise des Nervensystems zum Ziel hat. Laut Hartje grenzt der Begriff „kognitiv“ das Untersuchungsgebiet auf das Gehirn ein, speziell auf eine Art der Hirnforschung, die die Zusammenhänge zwischen Gehirn und Kognition untersucht. Unter Kognition wird dabei in der heutigen Verwendung eine ganze Reihe sehr verschiedener Funktionen verstanden: Wahrnehmung, Aufmerksamkeit, Sprache, Denken, Vorstellung, Gedächtnis, Handeln, Reagieren, Motorik und Emotion (Hartje 2012, 2).

Nach dieser Definition ist das Forschungsgebiet, mit dem sich kognitive Neurowissenschaftler beschäftigen können, also sehr breit und vielschichtig und erfordert unterschiedliche Sichtweisen auf das Gehirn und seine Arbeit. Dies macht es, wie wir in Abschnitt 4.2 sehen werden, nicht leicht, die Fragen und Ziele der kognitiven Neurowissenschaften nachzuzeichnen.

Der Psychologe Bradley Postle (Postle 2015, 4) identifiziert zwei Ebenen, auf denen der Begriff „kognitive Neurowissenschaften“ gebraucht wird: Auf einer breiten Ebene bezeichnet er die neurowissenschaftliche Erforschung der meisten Bereiche des menschlichen Verhaltens. In einem engeren Sinne bezeichnet der Begriff die Erforschung der neuronalen Grundlagen des Denkens, woraus es besteht, wovon es beeinflusst und kontrolliert wird. Mit solchen Fragen beschäftigen sich auch andere Disziplinen, wie Philosophie, Biologie, Anthropologie und weitere. Postle (Postle 2015, 2) grenzt die kognitiven Neurowissenschaften folgendermaßen von ihnen ab:

„What distinguishes cognitive neuroscience [...] from these and other disciplines, is its grounding in the methods and traditions of neuroscience, and the primacy that it places on understanding the biological bases of mental phenomena.“

Die kognitiven Neurowissenschaften weisen also einen nicht geringen Aspekt an kognitiver Psychologie auf. Diesen verbinden sie mit der Herangehensweise von Neurowissenschaftlern und ihren empirischen Forschungsmethoden. Und mit welchen Fragen genau befassen sie sich? Postle (Postle 2015, 3) räumt ein, dass es nicht leicht ist, die Grenzen der Disziplin kognitive Neurowissenschaften zu definieren. Durch die relativ frühe Verbindung neurowissenschaftlicher Methoden und der Erforschung menschlichen Verhaltens mit Bereichen, die von kognitiver Psychologie bearbeitet wurden (beispielsweise optische Wahrnehmung oder Gedächtnis), ist der Begriff „kognitive Neurowissenschaften“ älter als andere, eventuell explizitere Begriffe (wie beispielsweise affektive Neurowissenschaften). Dadurch wird der Begriff „kognitive Neurowissenschaften“ heute in vielen Kontexten verwendet, zu Postles Bedauern nicht immer mit der gleichen Bedeutung. Zusammengefasst kann nach Postle „kognitive Neurowissenschaften“ verwendet werden, um auf die Werkzeuge und Methoden zu verweisen, die für die Erforschung der neuronalen Basis menschlichen Verhaltens verwendet werden. Er nennt als Beispiele Gehirnschans, elektrische Messungen und Gehirnstimulationen. Diese Methoden werden wir in Abschnitt 4.3 eingehender betrachten.

Bei dieser Arbeit, erläutert Postle (Postle 2015, 2) weiter, bedienen sich kognitive Neurowissenschaftler Methoden, die auch Physiker und Biologen anwenden: Sie führen streng kontrollierte Experimente durch, die objektive, messbare Daten liefern, und versuchen, diese Daten mit mechanistischen Modellen darüber in Verbindung zu bringen, wie natürliche Systeme arbeiten.

Die Kernfrage, denen sie dabei nachgehen, definiert Postle (Postle 2015, 13) als:

„How does the functioning of the brain give rise to cognition and behavior?“

Anfang des 19. Jahrhunderts gab es zwei gegensätzliche Konzeptionen: *Localization of Functions* und *Mass Action*. Die Idee der *Localization of Functions* geht davon aus, dass es im Gehirn bestimmte Bereiche gibt, die sich auf bestimmte Aufgaben spezialisiert haben und für diese zuständig sind. Werden diese Hirnareale beschädigt, ist ihre Funktion beeinträchtigt, was sich auf das Verhalten der Patienten auswirkt. Ist beispielsweise das Sehzentrum gestört (zum Beispiel durch einen Schlaganfall), werden die Betroffenen eine irreversible Einschränkung ihres Sehvermögens erfahren. Wird dagegen das Hörzentrum beschädigt,

werden die Patienten schlechter hören, aber *ceteris paribus* weiterhin gut sehen können.

Die Idee der *Mass Action* dagegen sieht das Gehirn als eine große Masse, die nicht aus vielen Arealen mit unterschiedlichen Funktionen und Spezialisierungen besteht. Im Gegenteil arbeitet das Gehirn nach dieser Auffassung eher homogen, jeder Teil des Gehirns ist an allen möglichen kognitiven Funktionen beteiligt. Das Ausmaß der Schädigungen, die ein Patient erfährt, liegt demnach nicht an der lokalen Ausbreitung der Läsion, sondern einzig an der Menge des Gehirns, die geschädigt ist. Ist nur ein kleiner Teil des Gehirns betroffen, so wird die Beeinträchtigung der Fähigkeiten gering ausfallen, egal um welche Art es sich handelt. Wird allerdings ein großer Teil des Gehirns geschädigt, fällt auch die Beeinträchtigung größer aus.

Beide Konzeptionen waren in Teilen des 19. Jahrhunderts die jeweils vorherrschende Konzeption und befeuerten die Entwicklung der kognitiven Neurowissenschaften in die eine oder andere Richtung. Zuerst brachten Gall und Johann Caspar Spurzheim mit ihrer Lehre der Phrenologie die Idee der Lokalisation von Gehirnfunktionen zu größerer Verbreitung. Die Phrenologie, in den ersten beiden Jahrzehnten des 19. Jahrhunderts entwickelt, beruhte auf der Idee, dass das Gehirn in Bereiche aufgeteilt ist, die unterschiedliche Aufgaben erfüllen, beziehungsweise für bestimmte (Charakter-)Eigenschaften einer Person verantwortlich sind. Ist ein Gehirnbereich anatomisch besonders ausgeprägt, so ist es auch der zugehörige Teil des Körpers oder Charakters der Person. Gall und Spurzheim glaubten, dass Eigenschaften der Anatomie des Gehirns auch von außen, am Schädel einer Person, erkennbar sind. Ein stark ausgeprägter Teil des Gehirns zeigt sich durch eine Ausbuchtung des Schädels an jener Stelle, während ein wenig ausgeprägter oder gar kümmerlicher Bereich des Gehirns eine Einbuchtung im Schädel zur Folge hat. Sie erstellten eine Ordnung der verschiedenen „Geistes-Vermögen“ (Struve 1843, 25) und teilten sie in *Facultés affectives*, was Struve mit Sensitivität übersetzt, und *Facultés intellectuelles*, was Struve mit Intelligenz übersetzt, ein. Die *Facultés affectives* werden weiter unterteilt in Triebe und Gefühle, die *Facultés intellectuelles* in perzeptive und reflektive Vermögen (Struve 1843). In diese Kategorien werden dann die sogenannten phrenologischen Organe eingeordnet, wie beispielsweise Anhänglichkeit, Zerstörungstrieb und Nahrungstrieb, Sorglichkeit, Wohlwollen, Gewissenhaftigkeit und Beifallsiebe, Nachahmungstalent, Ordnungssinn und Witz, Gegenstandssinn, Tatsachensinn, Zahlensinn und noch viele weitere. Diese Organe bekommen in der Phrenologie ihren Platz am Gehirn und somit am Schädel zugewiesen, noch heute berühmt sind die Schädelmodelle mit aufgezeichneten Feldern für die Verortung der Organe. Die Ausprägungen der verschiedenen *Facultés* einer Person waren für den kundigen Phrenologen offen

erkennbar, nicht zuletzt mit Hilfe von Messwerkzeugen. Die Phrenologie verbreitete sich rasch und wurde auch nach Galls Tod im Jahr 1828 in Nordamerika und vor allem in England von Georg Combe weiter erforscht.

Inzwischen haben die kognitiven Neurowissenschaften herausgefunden, dass kleine, idiosynkratische Unterschiede in der groben Form des Gehirns einen sehr geringen bis gar keinen Einfluss auf den Charakter einer Person haben. Das war aber nicht der einzige Fehler der Phrenologen: Die Verortung der verschiedenen Organe im Gehirn, also die Zuweisung von Funktionen zu bestimmten Gehirnarealen, basierte nicht auf wissenschaftlichen Methoden und stellte sich inzwischen als komplett falsch heraus. Ebenso wenig wissenschaftlich oder systematisch gingen Gall und seine Nachfolger bei der Auswahl und Definition der Funktionen vor, die sie den Gehirnarealen zuschrieben. (Postle 2015, 8). Gall, Spurzheim und Combe verließen sich dabei vielmehr auf ihre eigene Intuition.

Doch so sehr sie mit ihrer Definition und Verortung von Gehirnfunktionen auch auf dem Holzweg gewesen waren, hatten die Phrenologen doch die Idee der Lokalisation von Funktionen zur Sprache gebracht und verbreitet. Postle (Postle 2015, 11) beschreibt, wie zu dieser Zeit, nicht zuletzt als Antwort auf die Verbreitung der Phrenologie, Ärzte und Wissenschaftler begannen, sich in der Erforschung der Lokalisation von Funktionen zu engagieren. Dies taten sie mit Hilfe falsifizierbarer Hypothesen, die sie formulierten und dann an kontrollierten, replizierbaren Laborexperimenten testeten. Weit verbreitet war dabei die Herangehensweise, die Veränderungen im Verhalten von Versuchstieren zu analysieren, die Schäden an bestimmten Strukturen im Gehirn zugefügt bekommen hatten. Diese Art der Forschung wurde später als Teil der Neuropsychologie bekannt. Sie versucht, das Funktionieren des Gehirns dadurch zu verstehen, dass systematisch Teile des Gehirns entfernt oder inaktiviert werden und dann beobachtet wird, wie sich das Entfernen jedes Teilstücks auf das Funktionieren des Gehirns, also auf das Verhalten und die Fähigkeiten der Tiere, auswirkt (Postle 2015, 15). Einer der bekanntesten und einflussreichsten Vertreter dieser Zeit war Pierre Flourens, der bereits in den 1820er Jahren neuropsychologische Experimente an Tauben und Hunden durchführte. Seine Ergebnisse widersprachen der Phrenologie vehement. Beispielsweise konnte er durch eine Schädigung des Cerebellum eine Beeinträchtigung des Gleichgewichtsvermögens der Versuchstiere feststellen, jedoch keine des Geschlechtstriebes, wie die Phrenologie vorausgesagt hätte. Doch Flourens kam nicht nur zu dem Schluss, dass die Verortung von Funktionen, wie sie die Phrenologie vorlegte, falsch sei, er konnte auch keinen Nachweis dafür erbringen, dass Funktionen überhaupt an bestimmten Stellen im Gehirn verortet sind. Seine Versuchstiere zeigten bei einer Schädigung des Cortex zwar unweigerlich eine Störung von Verhalten, das mit Ermessen, Erinnern oder

Perzeption zusammenhing. Doch diese Störungen beobachtete Flourens unabhängig davon, welcher Teil des Cortex geschädigt worden war. Daraus schloss er, dass alle Bereiche des Cortex gleichermaßen an diesen Fertigkeiten beteiligt waren. Zudem beobachtete Flourens, dass sich seine Versuchtauben, deren Cortex teilweise geschädigt worden war, oftmals wieder auf ein vor-experimentelles Verhaltensniveau erholten. Allerdings sah er keine Reparatur der geschädigten Materie vorliegen, weshalb er annahm, die intakten Teile des Gehirns hätten die zuvor gestörten Funktionen übernommen. Daraus entstand das Konzept der *Equipotentiality*, das besagt, dass jedes beliebige Stück Materie aus dem Cortex das Potential hat, jede Gehirnfunktion zu unterstützen. (Postle 2015, 11).

Flourens ging also davon aus, dass die phrenologische Idee der Lokalisation von Funktionen im Gehirn nicht zutreffend war, dass das Gehirn nicht in spezialisierte Areale aufgeteilt ist, sondern dass der gesamte Cortex (das cerebellum ist nach seinen Forschungen eine Ausnahme) an allen kognitiven Funktionen beteiligt ist. Flourens' Ergebnisse basierten, im Gegensatz zu phrenologischen Lehrmeinungen, auf den wissenschaftlichen Methoden seiner Zeit, was das Pendel in der wissenschaftlichen Gemeinschaft stark in Richtung *Mass Action* ausschlagen ließ. Sie blieb bis in die 1860er Jahre die vorherrschende Konzeption der Neuropsychologie.

Der Umschwung zurück zur *Localization-of-Functions*-Idee kam langsam, entwickelte sich dann aber zur goldenen Ära der *Behavioral Neurology*, die bis ins 20. Jahrhundert hinein andauerte. Sie begann mit einer Häufung von einzelnen Veröffentlichungen, in denen Forscher berichteten, ihre Untersuchungen deuteten darauf hin, dass Schäden an der linken Gehirnhälfte oder Schäden an vorderen Teilen des Gehirns eine Einschränkung der Sprachfähigkeiten von Patienten zur Folge hätten. Aus der Häufung dieser Berichte heraus entstand bald die Hypothese, vordere Bereiche der linken Hemisphäre des Gehirns seien wichtig für Sprache. (Postle 2015, 16).

Solche Überlegungen an sich waren bereits ein Schritt in Richtung Lokalisation von Funktionen. Ob sie allein der vorherrschenden *Mass-Action*-Konzeption hätten gefährlich werden können, ist eine andere Frage. Die Diskussion um die Lokalisation von Sprache nahm jedoch an Fahrt auf mit den Veröffentlichungen des französischen Chirurgen Broca. In den 1860er Jahren veröffentlichte er eine Reihe von Berichten über seine Fallstudien am Hospital von Bicêtre in Paris. Dort beobachtete er mehrere Patienten mit Sprachstörungen, die er nach deren Tod im Hospital auch obduzieren konnte. Der bekannteste seiner Fälle ist ein Mann namens Lelong, der als „Tan“ bekannt wurde, da dies die einzige Silbe war, die er äußern konnte. Lelong hatte seit seiner Kindheit an Epilepsie gelitten und mit 30 Jahren schwere Störungen in seiner Fähigkeit zu sprechen gezeigt, zehn

Jahre später begann zudem eine langsam fortschreitende rechtseitige Hemiparese. Einige Monate später behandelte Broca einen weiteren Patienten mit ähnlichen Symptomen der Sprachstörung wie Lelong. Über seine Beobachtungen an diesen beiden Patienten berichtete Broca erstmals im Jahr 1861 in den *Bulletins de la société anatomique*. Ihm war aufgefallen, dass es den Patienten nicht an Intelligenz fehlte, oder am Verständnis von Worten, sondern an der Fähigkeit, selbst Wörter zu produzieren. Er fand heraus, dass diese Unfähigkeit nicht an den motorischen Funktionen der Zunge oder des Sprechapparates liegen konnte, da die Patienten Zunge und Mund gut beherrschten, wenn sie Laute produzierten, die keine Worte waren. Vielmehr hatten die Patienten ihre Fähigkeit verloren, die Abfolge systematischer und koordinierter Bewegungen auszuführen, die die Silben hervorbringen, nach denen sie suchten. (Roth 2014, 13). Broca schreibt (Broca 1861, 333):

„Ce qui a péri en eux, ce n'est donc pas la faculté du langage, ce n'est pas la mémoire des mots, ce n'est pas non plus l'action des nerfs et des muscles de la phonation et de l'articulation, c'est autre chose, c'est une faculté particulière considérée par M. Bouillaud comme *la faculté de coordonner les mouvements propres au langage articulé*.“¹

Ohne diese *Faculté* ist nach Broca also keine Artikulation von Wörtern möglich. Er schließt aus seinen Beobachtungen, dass das Produzieren von Sprache klar verschieden ist vom Verständnis von Sprache, ähnlich wie bei einem Kind, das bereits viele Wörter verstehen kann, bevor es in der Lage ist, sie selbst zu verwenden. Die graduelle Perfektionierung der Sprachäußerungen bei Kindern beruht auf dem Aufbau eines Gedächtnisses, nicht für die Wörter an sich, sondern für die Bewegungen, die für die Artikulation der Wörter nötig sind. Und dieses Gedächtnis steht in keiner Verbindung zu den übrigen Erinnerungen oder der Intelligenz. Broca sieht seine Patienten in diesen Zustand vor dem Erlernen der Artikulation von Sprache zurückversetzt und nennt diesen Verlust der Sprachfähigkeit „aphémie“ (Broca 1861, 334, später setzte sich der Begriff Aphasie durch).

Bei der Dissektion von Lelongs Gehirn fand Broca eine etwa hühnereigroße Läsion in der vorderen linken Hemisphäre. Sie umfasste Teile des Frontallappens, der Insula, des *corpus striatum* und des vorderen oberen Temporallappens. Broca nahm an, dass die dritte vordere Hirnwindung wahrscheinlich für die Sprachfähigkeit entscheidend ist, da dieser Bereich früh von der Läsion betroffen zu sein

¹ Bei dem genannten M. Bouillaud handelt es sich um Jean-Baptiste Bouillaud, einem französischen Arzt, der kurz zuvor ebenfalls über die Lokalisation der Sprachfähigkeit in den vorderen Regionen des Gehirns veröffentlicht hatte.

schien und der Verlust der Sprache eines von Lelongs ersten Symptomen gewesen war. Die Dissektion des zweiten Patienten zeigte eine kleinere Läsion als die von Lelong, die vor allem die dritte untere vordere Hirnwindung betraf (heute Broca-Areal genannt). (Roth 2014, 14). Broca (Broca 1861, 356–357) hielt es aufgrund dieser beiden Ergebnisse für sehr wahrscheinlich, dass diese bestimmte Hirnwindung für die Fähigkeit zu artikulierter Sprache verantwortlich ist. Er räumte selbst ein, dass es für eine sichere Diagnose weitere Untersuchungen an mehr als seinen beiden Patienten geben müsste. Durch seine Ergebnisse sieht er jedoch Bouillauds Hypothese von der Existenz einer *Faculté* für die Koordination der Bewegungen für artikulierte Sprache bestätigt und sieht sie als ein wirksames Argument gegen die Annahme, dass die Sprachfähigkeit nicht an einem bestimmten Punkt im Gehirn verortet ist.

Brocas Hypothesen wurden nicht gleich von der wissenschaftlichen Gemeinschaft akzeptiert. Seine und ähnliche Ideen seiner Kollegen wurden noch jahrelang heiß diskutiert. Was Brocas Arbeit aber besonders machte, war seine für die damalige Zeit detaillierte Diskussion sowohl der Symptome der Patienten als auch der Anatomie des Gehirns in den untersuchten Regionen. Er beschränkte sich nicht auf eine grobe Beschreibung der vorderen linken Hemisphäre, sondern ging präzise auf eine bestimmte Region ein. Hinzu kommt, dass die Symptome seiner Patienten gerade den hervorstechendsten Aspekt der menschlichen Sprachfähigkeit betreffen, die Produktion von Sprache. Das macht den Verlust dieser Fähigkeit besonders interessant, nicht nur für die Forschung. Broca ging nicht nur direkt auf die wahrnehmbaren Symptome und Einschränkungen der Patienten ein. Er deckte auch auf, welche Aspekte der Sprachfähigkeit nicht betroffen waren und unterschied sie von Aspekten der Erinnerung und der Intelligenz. Schon bevor er die Dissektion der Gehirne vornehmen konnte, hatte er eine genaue Beschreibung der Krankheit und ihre möglichen Ursachen im Gehirn angestellt und eine Hypothese über die Ursachen der Einschränkungen aufgestellt. Mit dieser Vorgehensweise brachte Broca das erste (oder eines der) ersten klaren Beispiele für eine lokale Verortung einer genau umschriebenen Gehirnfunktion, die auf der Beobachtung der Patienten und anatomischer Forschung fußte. (Roth 2014, 14, 16). Neben diesen Aspekten der detaillierten und fundierten wissenschaftlichen Arbeit des Forschers Broca bringt Postle (Postle 2015, 17) einen weiteren zur Sprache, der Brocas Arbeit damals (und heute) viel Aufmerksamkeit erfahren ließ: Broca war es gelungen, eine Vorhersage zu bestätigen. Bouillaud und andere hatten vor Broca die Hypothese aufgestellt, eine Region in der linken vorderen Hemisphäre des Gehirns sei wichtig für die Sprachfähigkeit einer Person. Diese Hypothese konnte Broca eindeutig stützen. Gerade in einer Zeit, in der sich das Testen von Hypothesen als wichtiger Bestandteil wissenschaftlicher Arbeit eben etablierte,

war dies sicher ein beeindruckender Erfolg für die Befürworter der Lokalisation von Funktionen.

Der Durchbruch für die Lokalisationshypothese kam dann 1874 mit der Veröffentlichung Wernickes zum aphasischen Symptomenkomplex (Wernicke 1874). Waren nach Broca noch hitzige Debatten im Gange, konnte spätestens dieser Mediziner die meisten Bedenken auch der deutschen Kritiker ausräumen. Sie waren vor allem deshalb kritisch mit Broca gewesen, da zur damaligen Zeit in Deutschland, anders als in Frankreich, die Physiologie die vorherrschende Wissenschaft zur Erforschung von Hirnfunktionen war. Roth (Roth 2014, 21) führt dies auf den großen Einfluss von Hermann Helmholtz und Johannes Müller zurück und erläutert, was physiologische Modelle ausmacht:

„Physiological models emphasized the mapping of smaller subcomponents of functions onto physiological processes, and uncovering principles of biological function through experimentation and fine-scaled measurement.“

Broca hatte die Funktion zur Artikulation von Wörtern als eine eigene *Faculté* beschrieben und aufgrund einer klinisch-pathologischen Korrelation dafür eine Verortung im Gehirn definiert. Befunde, die auf klinischen Korrelationen und großen psychologischen Einheiten basierten, erschienen Physiologen allerdings weniger verlässlich und ließen sich nur schwer in physiologische Begriffe übersetzen. (Roth 2014, 21). Offenbar war ihnen Brocas Arbeit zu grob. Wernicke gelang in dieser Situation ein Brückenschlag. Er präsentierte ein neues Modell der Produktion von Sprache in einer Terminologie, die sich von Brocas unterschied und eher physiologischer Natur war. Wernicke war zu jener Zeit Assistenzarzt an der Irrenstation des Allerheiligenhospitals in Breslau. In dieser Tätigkeit behandelte er eine Reihe von Patienten, die Einschränkungen ihrer Hirnfunktionen aufwiesen, die damals oft als „Verwirrtheit“ (Wernicke 1874, 39) zusammengefasst wurden, sodass sie in Wernickes Abteilung eingeliefert wurden. Nach ihrem Tod konnte er einige der Patienten obduzieren, wodurch er in der Lage war, die Beeinträchtigungen seiner Patienten mit Hirnläsionen in Verbindung zu bringen. Hinzu kommen Fälle von Kollegen, die er in seine Veröffentlichung einbezieht.

Wernicke bestätigte Brocas Einschätzungen dahingehend, dass die von Broca beschriebene Aphasie durch eine Beschädigung der motorischen Komponente für Sprache verursacht wurde, die sich im vorderen *operculum* oder im Broca-Areal befindet. Dieses Areal beherbergt das Gedächtnis für das Wissen, wie Wörter produziert werden. Wernicke (Wernicke 1874, 16) entdeckte aus eigener Arbeit, und der von Kollegen, dass das Broca-Areal nicht das einzige Areal sein kann, das

als Sprachzentrum dient und kam so zur Definition einer weiteren, einer sensorischen Aphasie. Sie betrifft die sensorische Komponente für Sprache, die Wernicke im hinteren Teil der oberen temporalen Hirnwindungen verortete (später bekannt als das Wernicke-Areal). Dieses Areal beherbergt das Gedächtnis für das akustische Erscheinungsbild von Wörtern, also wie welche Wörter klingen. Bei der Broca-Aphasie können Patienten Wörter und Sätze, die man ihnen sagt, in der Regel gut verstehen. Sie können aber selbst keine Wörter formen, weil sie, grob gesagt, vergessen haben, wie das geht. Bei einer sensorischen Aphasie dagegen haben die Patienten vergessen, wie sich Wörter anhören. Sie erkennen Worte und Sätze nicht, die zu ihnen gesprochen werden. Sie selbst können aber Worte formulieren, ihre motorischen Fähigkeiten sind durch die sensorische Aphasie nicht eingeschränkt. Auch Geräusche, die keine Wörter sind, können sie identifizieren. Wernicke (Wernicke 1874, 40) beschreibt eine Patientin, die ihren eigenen Namen nicht erkennt und auf jeden Namen antwortet, bei dem man sie ruft. Vorgehaltene Gegenstände kann sie oft richtig benennen, aber sie versteht absolut nichts, was man ihr sagt. Das wird auch zum Problem, wenn die Patienten selbst Sätze formulieren. Im Grunde sind sie in der Lage dazu, sie besitzen einen großen Wortschatz und haben keine motorischen Einschränkungen (zumindest keine, die direkt mit der sensorischen Aphasie zusammenhängen). Jedoch hindert sie ihre eigene Unfähigkeit zum Verständnis von Sprache daran, ihre eigenen Sätze zu verstehen. Während sie selbst sprechen, verstehen sie sich nicht, sie können nicht überprüfen, was sie sagen. Das führt dazu, dass sie Wörter im Satz durcheinanderwerfen oder gar völlig neue, zum Teil unsinnige Wörter erfinden (Wernicke 1874, 42). Das macht es für Außenstehende schwer, die Patienten zu verstehen, während diese selbst ihr Sprachdefizit wenig bemerken.

Wernicke beschrieb in seinem Modell von der menschlichen Sprachfähigkeit also zwei verschiedene Sprachzentren im Gehirn. Aber auch deren Verbindung und die anderer sensorischer und motorischer Zentren untereinander erachtete er als unverzichtbare Komponente. Nach seiner Einschätzung laufen „Willensbewegungen“ (im Gegensatz zu „Reflexbewegungen“ (Wernicke 1874, 9)) folgendermaßen ab: Sinneseindrücke gelangen über die Sinnesorgane in das entsprechende sensorische Zentrum des Gehirns. Dort hinterlassen sie ein „Erinnerungsbild“ (Wernicke 1874, 10), das bei erneutem äußerem Eindruck als Reiz über Gehirnfasern in das zugehörige motorische Zentrum gelangt. Dort veranlasst diese Erregung eine Bewegung (in unserem Sprachbeispiel eine Äußerung). Auf unser Beispiel angewandt, dringen gesprochene Wörter an das Ohr, werden weitergeleitet an das Wernicke-Areal, das sensorische Zentrum für Sprache. Von dort wird der Reiz weitergeleitet an das Broca-Areal, wo eine motorische Ausgabe an die Muskeln des Mundraums und der Zunge veranlasst wird, um in

Wernickes Vokabular zu bleiben. Diesen gedanklichen Bogen, den der Reiz dabei beschreibt, nannte Wernicke psychischer Reflexbogen (Wernicke 1874, 19). Diese Idee vom Reflexbogen wurde später von Behavioristen zwar kritisiert. Allerdings lieferte er eine Anleitung zur Lokalisation von Aphasien. Denn eine Beschädigung an irgendeinem der Punkte zwischen dem Eindringen des Reizes in das Ohr und der Ausgabe von Reizen über die Muskeln kann zu einer Aphasie führen. Die motorische Aphasie, bei der das motorische Zentrum beschädigt ist, hatte Broca beschrieben, Wernicke ergänzte die sensorische Aphasie bedingt durch eine Beschädigung des sensorischen Zentrums und sagte das Auftreten einer weiteren möglichen Aphasie voraus, der Leitungsaphasie. Bei dieser Form ist die Verbindung zwischen sensorischem und motorischem Zentrum gestört. Die Folgen sieht Wernicke (Wernicke 1874, 26–27) darin, dass die Patienten alles verstehen, was man ihnen sagt und auch alles sprechen können. Allerdings haben die Reize aus dem sensorischen Zentrum nur wenig oder gar keinen Einfluss auf die Bildung der motorischen Reize im motorischen Zentrum. Daher fällt es Patienten schwer, die richtigen Wörter auszuwählen, ähnlich wie bei der sensorischen Aphasie. Da bei der Leitungsaphasie das sensorische Zentrum aber funktioniert, verstehen die Patienten ihre eigenen Äußerungen und bemerken, dass sie Wörter verwechseln oder falsch gebrauchen. Das ist oft mit einigem Unmut verbunden, die Patienten beginnen den Satz von Neuem und geraten wieder ins Stocken. Sagt man ihnen verschiedene möglicherweise passende Wörter vor, können sie die richtigen auswählen. Es sollte den Patienten aber möglich sein, das, was sie sagen möchten, vorher einzuüben. Weniger stark Betroffenen fällt es im Redefluss schwer, die richtigen Worte zu finden, sie sprechen langsam und stockend mit langen Pausen. Auch beim Lesen kann es zu Schwierigkeiten kommen, vor allem bei einzelnen Buchstaben. Wernicke (Wernicke 1874, 28–29) erklärt, dass beim Erlernen des Lesens das optische Bild eines Buchstabens mit einem Klangbild verbunden wird. Bei Patienten mit einer Leitungsaphasie wird diese Verbindung nicht mehr hergestellt, wodurch die einzelnen Buchstaben nicht mehr gelesen werden können. Interessanterweise sieht Wernicke dieses Problem aber nicht für ganze Wörter. Hat eine Person flüssig lesen gelernt, so liest sie nicht die einzelnen Buchstaben und setzt sie zu einem Wort zusammen, sondern begreift das ganze Wort auf einmal. In diesem Fall „[...] wird durch das geschriebene Wort ein bestimmter Begriff in ihm [dem Patienten] lebendig, er versteht das Geschriebene und findet dafür [...] wohl auch das richtige Wort.“ Später wurde diese Leitungsaphasie klinisch nachgewiesen (Roth 2014, 23).

Wernicke hatte einen anderen Ansatz, das Gehirn zu erklären als Broca. Er verband nicht nur eine einzelne Funktion mit einem bestimmten Ort im Gehirn, sondern gab einen differenzierteren Blick auf mögliche Symptome verschiedener

Arten von Sprachstörungen. Seine Herangehensweise war nicht nur geschätzter unter Physiologen, sondern wird auch als Ausgangspunkt für die konnektionistische Sichtweise des Gehirns als aus spezialisierten anatomischen Zentren, Verbindungen und verzweigten Systemen bestehend angesehen. Die aus Wernickes Schule erwachsenen Konnektionisten illustrierten ihre Modelle häufig mit Diagrammen, die die funktionalen Zentren im Gehirn und ihre Verbindungen aufzeigten. Das war eine neue Methode, die Informationsflüsse im Gehirn aufzuzeigen, und damit nicht nur Auswirkungen von Läsionen erklären zu können. Auch neue, überprüfbare Hypothesen und Vorhersagen über pathologisches Verhalten aufgrund von Läsionen konnten aufgestellt werden. Insgesamt entstand ein neues Verständnis über den Zusammenhang von Gehirn und Verhalten. (Roth 2014, 25, 46).

Im Anschluss an Wernicke entstand eine Reihe von weiteren konnektionistischen Arbeiten auch zu anderen Hirnfunktionen, wie beispielsweise Hugo Karl Liepmanns Untersuchungen zu Apraxien, einer Reihe motorischer Störungen, oder Heinrich Lissauers Forschungen mit visueller Agnosie, bei der Betroffene Gegenstände und Gesichter zwar sehen, aber nicht erkennen können. Doch bereits um die Jahrhundertwende wurde immer stärkere Kritik am Konnektionismus laut, die sowohl seine neurologische als auch seine psychologische Seite betraf. Neurologe Henry Head bezeichnete die Konnektionisten als „diagram makers“ (Roth 2014, 47, zitierend Head 1926, 63), die ihre Fälle so zurechtbogen, dass sie in ihren Rahmen passten. Head war der Meinung, die Konzepte der Konnektionisten, ob motorische oder sensorische, seien zu weit gefasst, um Erklärungen liefern zu können. Roth (Roth 2014, 48) sieht den damals vorhandenen Einfluss von Head in seiner Disziplin als Grund für ein schwächer werdendes Interesse der Neurologie an konnektionistischen Ideen. Ein ähnlicher Zugpferdeffekt scheint sich auch in psychologischen Kreisen abgespielt zu haben. Sigmund Freud war kein Freund der Konnektionisten und kritisierte ihre Hypothesen als nicht zutreffend. Seine Erfahrung mit Patienten widersprach den Schlussfolgerungen von Forschern wie Ludwig Lichtheim (Roth 2014, 28). Auch Vertreter der damals populären Gestaltpsychologie argumentierten gegen konnektionistische Modelle. Ihr Augenmerk lag nicht auf der Identifikation von Funktionszentren und ihrer Verbindungen, sondern auf der Organisation des Verstandes im Sinne von Mustererkennung (Roth 2014, 48).

Vor allem aus Amerika kam Kritik von Vertretern eines radikalen Behaviorismus, der sich rasch ausbreitete. Die Behavioristen verneinten die Verlässlichkeit subjektiver Berichte. Sie sahen das Gehirn als Black Box, weshalb sie sich auf die Messung von Input und Output konzentrierten. Einzig das physikalische Verhalten und Gehirn von Probanden ist allgemein beobachtbar und daher als einzige

Informationsquelle gültig. Hypothesen darüber, was sich im Gehirn abspielt, hielten die Behavioristen für wissenschaftlich unzulässig, ebenso wie Hypothesen zum Bewusstsein. Studien wie die der konditionierten Hunde des Physiologen Iwan Pavlov bestärkten auch unter Psychologen die Ansicht, dass letztlich jegliches Verhalten anhand einfacher, reflexbasierter Verhaltenselemente erklärbar sei (Baars et al. 2010, 22). Infolgedessen treten Bewusstseinsstudien und die Lokalisierungshypothese im Laufe des 20. Jahrhunderts sowohl in der Psychologie als auch in der Physiologie in den Hintergrund. Die Arbeit von Wernicke und anderer Konnektionisten galt als fehlgeleitet, unwissenschaftlich und irrelevant (Roth 2014, 49).

Die Dominanz des Behaviorismus beginnt erst in den 1960er Jahren aufzuweichen, als der ursprünglich behavioral ausgebildete Neurologe Norman Geschwind die Literatur der Forschung um die Jahrhundertwende wieder aufgreift. Er sieht die Vorurteile nicht bestätigt, sondern findet detaillierte Analysen und von Verständnis zeugende Schlussfolgerungen. Vor allem Läsionsstudien prägen daraufhin seine eigene Forschung. Seine neue Begeisterung für den Konnektionismus teilt er in Veröffentlichungen (vor allem Geschwind 1965), was nicht nur den Konnektionismus, sondern auch die Lokalisationsidee sowie die Bewusstseinsforschung im Allgemeinen wieder in das Bewusstsein der Neurologen und Psychologen brachte. Allerdings dauerte es noch weitere zehn bis zwanzig Jahre, bis die Enttäuschung der Forscher mit dem radikalen Behaviorismus soweit gediehen war, dass sie in ihrer Arbeit neue Wege einschlugen.

4.2 Ziele der kognitiven Neurowissenschaften

Im vorangegangenen Abschnitt 4.1 haben wir einen kurzen Überblick darüber gewonnen, welche Fragen die kognitiven Neurowissenschaften und ihre Vorgänger beschäftigen und welche Konzeptionen nicht nur in der Geschichte von großer Bedeutung waren, sondern auch heute noch die Forschung anleiten und zu neuen Fragen inspirieren. Nun werfen wir in diesem Abschnitt 4.2 einen Blick auf die gegenwärtige Arbeit der kognitiven Neurowissenschaftler. Was wollen sie mit ihrer Forschung erreichen, nach welcher Art von Modellen suchen sie? Wir werden feststellen, dass sich die Modelle stark unterscheiden im Hinblick auf das Gebiet der kognitiven Neurowissenschaften, das sie beschreiben, sowie auf ihre Abstraktion und ihren Auflösungsgrad, also auf welcher Ebene sie das Gehirn und seine Arbeit untersuchen – von einzelnen Synapsen bis zum gesamten Gehirn. Das liegt nicht zuletzt an der Vielseitigkeit der kognitiven Neurowissenschaften, die einen generellen Umriss mit einem universellen Modell schwierig

macht. Um zu verstehen, welche Konzeption heute, nach wechselvollen zwei Jahrhunderten, die vorherrschende ist und welche Ziele sich die Forscher unterschiedlicher Forschungsschwerpunkte setzen, werden wir uns nach einer kurzen Einführung zwei Forschungsfelder ansehen, Gedächtnis und Emotionen, sowie zwei aktuelle Großprojekte, das Human Brain Project und die BRAIN Initiative, mit je ihren Fragen und Zielen. Unter den unzähligen neurowissenschaftlichen Forschungsprojekten lässt sich vermutlich an solch umfangreichen am ehesten erkennen, welche Modelle die modernen Neurowissenschaften suchen und welche die heutigen großen Fragen sind.

Nachdem der Behaviorismus für große Strecken des 20. Jahrhunderts die vorherrschende Konzeption in Psychologie und Physiologie war, begann sich in den 1970er Jahren eine gewisse Enttäuschung unter den Forschern breit zu machen. Sicherlich haben Behavioristen wie Burrhus Frederic Skinner mit ihrer Arbeit neue Erkenntnisse hervorgebracht. Doch bestimmte Fragen blieben mit behavioristischen Mitteln weiterhin ungeklärt. Antworten auf Fragen beispielsweise zu Gedächtnis, Vorstellung, Unterbewusstsein und generell des Bewusstseins wurden innerhalb der Black Box vermutet, die das menschliche Gehirn im Behaviorismus darstellte. Trotz aller propagierter Unwissenschaftlichkeit von Bewusstseinsfragen und unbeobachtbarer Abläufe wollten einige Forscher doch die Black Box öffnen. Sie waren nicht der Meinung, dass Bewusstsein ohne Funktion oder bewusste Entscheidungen eine Illusion seien. Heute beschäftigen sich kognitive Psychologen wieder mit Bewusstsein, Unterbewusstsein, Gedächtnis, Vorstellung und vielem mehr. Beim Lesen von Texten zur Geschichte der Kognitionswissenschaften bekommt man beinahe das Gefühl, es herrsche eine Stimmung von Befreiung. Der Psychologe Allan Hobson (Hobson 2009, xi) fasst die Situation in bildhaften Worten zusammen:

„Suddenly it is spring. We have survived the long winter of behaviourism. We have tripped over the traces of reflexology. We are about to walk out the long shadow of psychoanalysis. This, surely, is cause for celebration. Consciousness, like sleep, is of the Brain, by the Brain, and for the Brain. A new day is dawning.“

Sicherlich, meint auch Hobson (Hobson 2009, xii), sind die Methoden des Behaviorismus nicht ganz vergessen, aber sie sind nun nicht mehr der begrenzende Horizont der Forschung. Ein Grund für den neuen Enthusiasmus für die Vorgänge im Gehirn ist sicherlich auch die Erfindung und starke technische Weiterentwicklung von bildgebenden Verfahren (s. Abschn. 4.3). Längst sind vor allem in Amerika nicht mehr nur medizinische Zentren mit fMRTs oder EEGs ausgestattet, doch noch vor dreißig Jahren war die Werteausgabe einer EEG-Untersuchung

eine Nadellinie auf einer Papierrolle. Durch den Einsatz von Computertechnik wurde ein großer Schritt zur Verfügbarkeit und Vereinfachung von kognitiven Experimenten getan. Diese neue Quelle für Informationen und Daten schafft neue Möglichkeiten für Forschung, die vielseitig genutzt werden.

Zu Beginn dieses Kapitels haben wir gesehen, wie viele Begriffe unter das neurowissenschaftliche Verständnis von Kognition fallen: Wahrnehmung, Aufmerksamkeit, Sprache, Denken, Vorstellung, Gedächtnis, Handeln, Reagieren, Motorik. Zu allen diesen Gebieten gibt es Forschungen, was eine Vorstellung davon gibt, wie vielseitig die Arbeiten der Neurowissenschaftler sind und welch breites Feld sie bearbeiten. Zusätzlich gibt es unterschiedliche Ebenen der Untersuchung. Manche Fragen beschäftigen sich mit der Funktionsweise von Synapsen, oder auch nur einzelner Botenstoffe. Andere untersuchen axonale Verbindungen zwischen Neuronen und wieder andere die Lokalisation von Funktionen in ganzen Verbänden von Neuronen. Dann gibt es Fragen zur Arbeit des gesamten Gehirns, oder zum gesamten Prozess von der Reizaufnahme bis zur Reaktionsausgabe. Auf all diesen Ebenen werden die verschiedenen Aspekte von Kognition untersucht. Das bringt eine diverse Priorisierung von Fragen mit sich. Das heißt, welche Frage genau die große, zentrale, spannende Frage der kognitiven Neurowissenschaften ist, sieht jeder Forschende anders, je nachdem, aus welchem Forschungsfeld er stammt. Das macht eine Abgrenzung der einzelnen Aspekte von Kognition, wie Gedächtnis oder Vorstellung, schwierig. Kognitive Neurowissenschaftler suchen nicht nach dem einen einheitlichen Modell, das die Kognitionsleistungen des Gehirns beschreiben könnte, weil das, zumindest im Moment, nicht möglich ist. Vielmehr werden auf den verschiedenen Ebenen der Forschung eigene Modelle erstellt, deren Qualität und Stärke durch die verschiedenen Reifestufen der einzelnen Gebiete bedingt sind.

Was sie aber alle eint, ist die wissenschaftliche Vorgehensweise, falsifizierbare Hypothesen aufzustellen und diese in Experimenten zu prüfen. Beispiele für Forschungsfragen und Experimente, wie sie in den kognitiven Neurowissenschaften auch für die neuroökonomische Arbeit Anwendung finden, werden wir in Kapitel 5 betrachten.

Aufgrund der Breite des Forschungsgebietes und der Tiefe der möglichen Ebenen lässt sich ein einheitliches Ziel aller kognitiven Neurowissenschaftler also nicht beschreiben. Wir können aber einzelne Forschungsfelder betrachten, die in den kognitiven Neurowissenschaften derzeit bearbeitet werden, und auf diese Weise neurowissenschaftliche Fragen und Ziele nachvollziehen. Die Anzahl der Bearbeitungsgebiete ist immens, an dieser Stelle können wir nicht alles abdecken. Daher beschränken wir uns auf zwei Forschungsfelder, die auch für die neuroökonomische Forschung relevant sind. Dabei geht es um Gedächtnis, das

eng mit Lernprozessen zusammenhängt, und um Emotionen. In Kapitel 5 werden wir neuroökonomische Experimente betrachten, die sich mit diesen Themen beschäftigen.

4.2.1 Gedächtnis

Das Gedächtnis ist kein einheitliches Objekt, für das Gedächtnis und das mit ihm verknüpfte Lernen sind mehrere Systeme zuständig. Die Aufgabe der Funktion Gedächtnis ist nicht so nebensächlich, wie man meinen könnte, sondern erfüllt elementare Voraussetzungen für das tägliche (Über-)Leben eines Individuums. Auch beim Gedächtnis werden bewusste von unbewussten Prozessen unterschieden. Postle (Postle 2015, 306) betont die Wichtigkeit der Gedächtnisleistungen für unser Leben und unsere Identität:

„At the conscious level, it [memory, Anm. von mir] creates our sense of identity, and provides the soothing, subjectively continuous internal narrative that enables us to go about our business without constantly stopping to ask 'Why am I here? What was I doing 5 minutes ago?' At an unconscious level, it shapes our habits, our proclivities and dispositions, and makes us who we are.“

Nicht ohne Grund fühlen sich Patienten mit retrograder Amnesie ihrer Identität beraubt, während Patienten mit anterograder Amnesie² häufig berichten, sie fühlten sich, als seien sie gerade aus einem Traum erwacht und hätten keine Ahnung, wo sie sind, wie sie an diesen Ort kommen, oder was sie hier tun (Postle 2015, 292). Ohne Gedächtnis gibt es also kein eigenes Identitätsbewusstsein und auch nicht Postles *Internal Narrative*, das uns den Weg weist. Ohne Gedächtnis ist auch das Lernen nicht möglich, denn zumindest Fakten und Ereignisse, die eines bewussten Wiederaufrufens bedürfen, sind dann verloren und können nicht als Referenz für ähnliche zukünftige Ereignisse dienen.

Neuroökonomien interessieren sich für Gedächtnis und Lernen, weil es unsere Art, Entscheidungen zu treffen, beeinflusst. In einem ökonomischen Spiel, das

² Bei einer retrograden Amnesie verlieren die Patienten ihr Gedächtnis für die bereits zurückliegende Zeit, während Patienten einer anterograden Amnesie ihr Gedächtnis für die vorausliegende Zeit verlieren, relativ zum Zeitpunkt des Gedächtnisverlusts. Erleidet also ein Patient an Tag X eine retrograde Amnesie, betrifft der Gedächtnisverlust die Zeit seines Lebens bis ungefähr zu diesem Tag X. Bei einer anterograden Amnesie bleibt diese Zeit vom Gedächtnisverlust verschont, jedoch verliert der Patient von Tag X an in unregelmäßigen Abständen immer wieder aufs Neue sein Gedächtnis für die Zeit seit Tag X. Am Beispiel von Henry Molaison betrachten wir das genauer (s. S. 70).

über mehrere Runden gespielt wird, sind die Spieler nach Runde 100 erfahrener als bei Runde 1. Die Informationen über das eigene Verhalten, das Verhalten der Mitspieler und die Auswirkungen der jeweiligen Handlungen werden im Gedächtnis gespeichert. Sie dienen bei weiteren Spielrunden als Entscheidungshilfe, sei es für eine *Tit-for-Tat*-Strategie oder um beispielsweise auf die sozialen Präferenzen des anderen Spielers zu reagieren. Nehmen wir als Beispiel das Ultimatumspiel: Spieler 1, der das Angebot an Spieler 2 abgibt, wird im Laufe mehrerer Spielrunden Informationen darüber sammeln, welche Angebote Spieler 2 annimmt oder ablehnt. Als erfolgreicher Spieler sollte Spieler 1 aus diesen Informationen ableiten, welche Angebote er in Zukunft machen sollte, um am Ende nicht leer auszugehen. Spieler 2 wiederum sammelt Informationen darüber, welche Angebote Spieler 1 macht und welche Reaktion das eigene Ablehnen oder Annehmen der Angebote in der nächsten Runde zur Folge hat. Macht Spieler 1 in Runde 2 ein besseres Angebot, wenn Spieler 2 das Angebot aus Runde 1 abgelehnt hat? Lassen wir die beiden Spieler nach 10 Runden die Rollen tauschen. Vielleicht wird ein fair bedachter Spieler 2 aus den ersten 10 Spielrunden nun auch gleich ein faires Angebot an den Mitspieler machen. Dazu braucht es wieder die Informationen aus dem Gedächtnis, die im Laufe der kommenden Spielrunden auch als Grundlage für das Erlernen einer eigenen Angebotsstrategie dienen werden. Gedächtnis- und Lernleistungen werden auch bei Spielen wie beispielsweise dem *Trust Game* oder den Experimenten von Smith aus Kapitel 3 abgefragt. Gedächtnis und Lernleistung sind also wichtige Faktoren beim Treffen von Entscheidungen. Wer schneller lernt, könnte im Vorteil sein und auch, woran sich die Spieler genau erinnern, ist wichtig. Erinnert sich Spieler 2 in der 11. Runde des Ultimatumspiels vor allem an die großzügigen Angebote von Spieler 1 aus den vorherigen Runden, könnte Spieler 2 selbst gute Angebote machen. Sind jedoch vor allem die niedrigen Angebote im Gedächtnis geblieben, könnte Spieler 2 nun vor allem niedrige Angebote vorlegen.

Derzeit sind Gedächtnis und Lernen Gegenstand intensiver neurowissenschaftlicher Forschung, die diverse und nicht nur kompatible Antworten hervorbringt. Im Laufe der vergangenen Jahre haben sich dabei auch neue Fragen aufgetan, die sich zu eigenen Forschungsfeldern entwickelten: In den 1950er Jahren machte die Gedächtnisforschung einen großen Sprung mit dem Fall des Patienten Henry Molaison aus den USA. Er hatte im Alter von fünfzehn Jahren einen ersten epileptischen Grand-Mal-Anfall³. In den darauffolgenden Jahren verschlechterte

³ Ein Grand-Mal-Anfall, auch tonisch-klonischer Anfall genannt, ist eine starke Form von epileptischen Anfällen, bei denen es in der Regel zu Bewusstlosigkeit und Krampfanfällen kommt.

sich sein Zustand stark, sodass er mit 27 Jahren das Haus nicht mehr verlassen konnte. Nachdem alle anderen Therapien ausgeschöpft waren, ohne dass sich eine Besserung gezeigt hätte, schlug ihm ein Neurochirurg die Entfernung seines Hippocampus vor, da dieser Teil des Gehirns damals als das am stärksten epileptogene Gewebe im Gehirn bekannt war. Molaison willigte ein und tatsächlich besserte sich seine Epilepsie erheblich, jedoch verlor er die Fähigkeit, sich für den Rest seines Lebens bewusst an irgendetwas zu erinnern, was er von dieser Zeit an bis zu seinem Tod erlebt hat. (Postle 2015, 288).

Molaison litt nach dieser Operation an anterograder Amnesie. Zwar erlitt er ebenfalls eine retrograde Amnesie, allerdings beschränkte diese sich auf wenige Jahre vor der Operation. Er wusste also, wer er war und erkannte die Personen, Orte und Abläufe, die er auch bis zu diesen paar Jahren vor der Operation gekannt hatte. Sein Problem bestand darin, dass er alle Ereignisse und Fakten, die er vom Tag der Operation an aufnahm, in unregelmäßigen Abständen dauerhaft vergaß. Dabei ging es ausschließlich um Ereignisse und Fakten, die bewusst wieder hervorgeholt werden mussten. Dagegen war Molaison in der Lage, neue Fertigkeiten und Gewohnheiten zu erlernen, die keiner bewussten Erinnerung bedürfen. (Postle 2015, 289). Dabei hatte er immer wieder das Gefühl, gerade erst aus einem Traum erwacht zu sein, sodass er sich nicht erinnern konnte, was er gerade getan hatte, oder wohin er gehen wollte. Er musste sich in seiner Umwelt ständig neu orientieren. Dieser Vergleich mit dem Erwachen aus einem Traum wurde auch von anderen Patienten mit anterograder Amnesie gezogen. Ein weiterer bekannter Patient ist der britische Musikwissenschaftler Clive Wearing, dessen Hippocampus durch eine Herpes Enzephalitis geschädigt worden war. Er beschrieb sich als ständig verwirrt und wie gerade aus einem Traum erwacht. Er führte über seine Gedanken Tagebuch, in das er immer wieder dieselben Eintragungen schrieb: „10:49 am I am totally awake – first time“, und direkt darunter, „11:05 am I am perfectly awake – first time“, und so weiter. Immer im Abstand von 5 bis 45 Minuten vergisst er alles Vorgegangene und fühlt sich wieder wie gerade erwacht. (Postle 2015, 292).

Die Erkenntnisse aus der Arbeit mit Patienten wie Molaison und Wearing gaben ihren Ärzten nicht nur Hinweise darauf, wie der Hippocampus in die Gedächtnisleistung des Gehirns eingebunden ist. Sie erlaubten auch die Skizzierung der Funktion Gedächtnis und ihrer Prozesse. Grob wird zwischen Kurzzeitgedächtnis/Arbeitsgedächtnis und Langzeitgedächtnis unterschieden, die unterschiedlich strukturiert werden. Im Gegensatz zum Langzeitgedächtnis erfüllt das Arbeitsgedächtnis die Aufgabe, Informationen kurzzeitig zu speichern und zu verarbeiten. Das gängigste Modell des Arbeitsgedächtnis, das modulare Arbeitsgedächtnismodell, stammt von Alan Baddeley und Graham Hitch aus den

1970er Jahren. Es unterscheidet in seiner Grundform drei Haupteinheiten, die für das Kurzzeitgedächtnis zusammenarbeiten (Hart 2016, 149): Die zentrale Exekutive (*Central Executive*) und die beiden Untersysteme visuell-räumlicher Notizblock (*Visuospatial Sketchpad*) und die phonologische Schleife (*Phonological Loop*).⁴ Der visuell-räumliche Notizblock verarbeitet eingehende visuelle Wahrnehmungen, während die phonologische Schleife akustische (auch sprachbasierte) Informationen verarbeitet. Beide Untersysteme haben die Aufgabe, die eingehenden Informationen, die im Kurzzeitgedächtnis aufbewahrt werden sollen, immer wieder zu wiederholen (*Rehearsal*), um den Inhalt zugänglich zu erhalten und nicht zu verlieren. Dazu werden sie auch in einem Speicher abgelegt. Allerdings nur für die kurze Zeit von wenigen Sekunden, dann ist der Inhalt meist verloren. Die zentrale Exekutive kontrolliert die beiden Untersysteme und greift bei Bedarf auf die gespeicherten Informationen zurück. Ein gängiges Beispiel ist eine Situation, in der man am Telefon beispielsweise eine andere Telefonnummer diktiert bekommt und sie dann im Kurzzeitgedächtnis behalten muss, bis Stift und Papier bei der Hand sind, oder man die Nummer nach Beenden des Gesprächs gleich in das Telefon eingibt. In diesem Fall wird im modularen Arbeitsgedächtnismodell die Telefonnummer von der phonologischen Schleife empfangen und immer wiederholt, damit sie nicht vergessen wird. Wenn Stift und Papier oder das Telefon parat sind, greift die zentrale Exekutive auf die gespeicherte Nummer zurück und leitet die Informationen an die an der folgenden Handlung beteiligten Module weiter. Je besser das Arbeitsgedächtnis Informationen verarbeiten kann, desto länger oder genauer stehen sie dem Individuum zur Verfügung. Ein häufig genutzter Test zur Messung der Leistungsfähigkeit des Arbeitsgedächtnisses von Personen ist der *Digit Span Test*. Bei diesem Test werden der Versuchsperson Ziffern vorgelesen mit einer Geschwindigkeit von einer Ziffer pro Sekunde. Am Ende einer Ziffernreihe werden die Versuchspersonen gebeten, die Ziffern vorwärts zu wiederholen. Die längste Ziffernreihe, die die Person fehlerfrei wiederholen kann, ist ihre *Forward Span* und gilt als Maß für ihre Fähigkeit, Informationen zu erfassen. In einem zweiten Teil des *Digit Span Test* bekommen die Versuchspersonen wieder Ziffern vorgelesen, diesmal sollen sie die Reihe jedoch rückwärts wiedergeben. Die längste Reihe in diesem Durchgang ist die *Backward Span* der Versuchsperson und gibt Auskunft über ihre

⁴ Im Laufe der Jahre wurde das Modell um ein weiteres Untersystem, den Episodischen Puffer (*Episodic Buffer*) erweitert. Für unsere Zwecke ist das einfache Modell allerdings ausreichend, da es leicht verständlich ist und die Idee hinter dem Modell zeigt. Für weitere Informationen zum modularen Arbeitsmodell und neueren Modellen des Arbeitsgedächtnis siehe Karnath 2012.

Fähigkeit, Informationen zu verarbeiten. (Hart 2016, 149) Denn in diesem zweiten Teil des Tests müssen die Informationen nicht nur gespeichert und wieder zur Verfügung gestellt werden, sondern auch manipuliert. Die erste Ziffer, die schon am längsten im Speicher liegt, sollte auch zuerst den Speicher wieder verlassen, um eine zu lange Liegedauer zu verhindern, während derer sie verloren gehen könnte. In diesem Fall allerdings sind die ersten Ziffern nicht nur die ersten, die im Speicher abgelegt werden, sondern auch die letzten, die ihn wieder verlassen.

Versuche mit fMRT-Geräten legen nahe, dass bei den Aufgaben des Arbeitsgedächtnisses der dorsolaterale präfrontale Kortex eine Schlüsselrolle spielt. Zusätzliche Läsionsstudien lassen die Beteiligung von Hirnregionen wie dem oberen Peritallappen oder dem ventrolateralen präfrontalen Kortex nicht ausschließen. Die genauen Zusammenhänge sind noch Gegenstand der aktuellen Forschung.⁵

Auch die Forschungen darüber, wie genau der Hippocampus das Langzeitgedächtnis beeinflusst oder welche anderen Hirnregionen außerdem welche Aufgaben genau erfüllen, sind noch nicht abgeschlossen. Beim Langzeitgedächtnis werden zwei Formen unterschieden: Das deklarative und das nondeklarative Gedächtnis. Beide sind für unterschiedliche Bereiche zuständig. Das deklarative Gedächtnis, auch *Explicit Memory* genannt, kommt bei Aufzeichnungen von Fakten und Ereignissen zum Einsatz, die bewusst zugänglich sind. Im Gegensatz dazu umfasst das nondeklarative Gedächtnis, auch *Implicit Memory* genannt, jedes Beispiel erfahrungsbasierter Plastizität (diesen Begriff werden wir weiter unten genauer betrachten), die das Verhalten eines Individuums auf eine Art beeinflusst, die nicht notwendigerweise eine bewusste Zugänglichkeit verlangt, um zum Ausdruck zu kommen. Darunter fallen beispielsweise Konditionierung oder motorisches Lernen. (Postle 2015, 290). Wenn ich also meine Heimatadresse oder den Geburtstag meiner Mutter aufsagen soll, benötige ich dafür das deklarative Gedächtnis, da diese Informationen in dem Moment bewusst abgerufen werden. Wenn aber beispielsweise ein Pianist ein schwieriges Stück mit vielen Läufen einübt, wird sein Gehirn die Fingerbewegungen, die für die schnellen Tonabfolgen benötigt werden, im nondeklarativen Gedächtnis ablegen. Denn sie werden nicht bei jedem Spielen bewusst abgerufen, sondern sind Teil des motorischen Lernens, sodass die einzelnen Finger ab einem gewissen Übungsstand nicht mehr einzeln bewusst an die richtigen Stellen bewegt werden müssen. Einmal angeschoben, wird der Vorgang wiedererkannt und automatisch wiedergegeben.

Beide Seiten des Langzeitgedächtnis liefern schematisch die gleichen Prozessschritte: Ihre Aufgaben sind *Encoding*, *Storage* und *Retrieval* von Informationen,

⁵ Für weiterführende Informationen siehe beispielsweise Hart 2016, 149–155.

wobei *Encoding* bedeutet, dass die eingehenden Informationen so verarbeitet werden, dass sie später wieder erinnert werden können. Alle drei Prozesse funktionieren über unterschiedliche Signale, die über unterschiedliche Synapsentätigkeiten zwischen den Neuronen weitergegeben werden. Glimcher fasst das zusammen (Glimcher 2014b, 74):

„The biochemical mechanism by which information is stored in the nervous system over periods of days or longer is a process of synaptic modification. If one thinks of the passage of an action potential from one neuron to another as a transfer function, memories are encoded through changes in those transfer functions.“

Was Glimcher hier erläutert, beruht auf dem Prinzip der Hebbschen Plastizität, nach dem Psychologen Donald Hebb benannt, der sie 1949 als Erster beschrieb (Hebb 1949). Hebbs Hypothese (Hebb 1949, 62) lautete:

„When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.“

Die Stärke der synaptischen Übertragung kann sich also verändern. In den 1960er und 1970er Jahren konnte diese Hypothese auch mit empirischer Evidenz unterlegt werden. Seither werden mehrere Arten von Aktivitäten unterschieden. Es wurde entdeckt, dass Synapsen nicht nur kurzzeitige Aktivierungen übertragen können, sondern auch Langzeit-Potenzierungen (*long-term Potentiation, LTP*) auftreten, ebenso Langzeit-Depressionen (*long-term Depression, LTD*). Bei beiden Effekten wird die Synapse zwischen zwei Zellen im Vergleich zum kurzen Aktionspotenzial gestärkt. Werfen wir einen kurzen Blick auf diesen Plastizitätsprozess, ohne auf die biochemischen Details einzugehen: Postle (Postle 2015, 295–298) beschreibt NMDA-Rezeptoren (eine Form der Glutamatrezeptoren an Zellmembranen) als anschauliches Beispiel, wie Plastizität und Erinnerung zusammenhängen. NMDA-Rezeptoren haben die Eigenschaft, nur dann aktiviert werden zu können, wenn zwei Ereignisse gemeinsam auftreten. Erstens die Freisetzung von Glutamat durch das präsynaptische Terminal der Synapse, die gestärkt werden soll, und zweitens eine vorangegangene Depolarisation des postsynaptischen Terminals. Also nur bei einer vorangegangenen Aktivierung der postsynaptischen Zelle kann der NMDA-Rezeptor aktiviert werden. Wie dies mit dem Gedächtnis zusammenhängt, erläutert Postle anhand der Pavlovschen Hunde: Wird den Hunden Futter vorgesetzt, werden olfaktorische Neuronen der Tiere aktiviert, was wiederum Neuronen aktiviert, die für die Verdauung verantwortlich sind. Dies liefert die benötigte Depolarisation für die NMDA-Rezeptoren.

Wenn nun während dieser Zeit zusätzlich die berühmte Glocke geläutet wird, setzen auditive Neuronen Glutamat in den synaptischen Spalt frei. Darunter sind auch Synapsen, die sie mit digestiven Neuronen verbinden. Dort aktiviert das Glutamat die NMDA-Rezeptoren. Diese Aktivierung, vor allem wiederholt, kann Langzeit-Potenzierungen auslösen. Das heißt, beim zukünftigen Erklingen der Glocke werden die nun gestärkten Synapsen zwischen auditiven und digestiven Neuronen feuern und so in den Hunden die Erinnerung an Futter hervorrufen.

Die genauen Abläufe bei Langzeitpotenzierungen sind Gegenstand der aktuellen Gedächtnisforschung. Es bleiben noch Fragen nach der genauen notwendigen chemischen Zusammensetzung, um LTPs auszulösen, oder welche prä- und postsynaptischen Veränderungen ihnen zugrunde liegen und viele weitere.⁶

4.2.2 Emotionen

Der zweite Forschungsbereich der kognitiven Neurowissenschaften, den wir betrachten wollen, sind Emotionen. Die Erforschung der neuronalen Basis von Emotionen erstreckt sich über ein breites Feld an Fragen und wird von einer eigenen Untergruppe der kognitiven Neurowissenschaften bearbeitet, den affektiven Neurowissenschaften. Auch in diesem Bereich wurde in den vergangenen dreißig Jahren, vor allem durch die Entwicklung der bildgebenden Instrumente, einiger Fortschritt erzielt. Allerdings sind auch hier noch viele Fragen offen und neue Erkenntnisse werfen wieder neue Fragen auf. Die Kontroverse beginnt bereits beim neurowissenschaftlichen Verständnis von Emotionen. Postle (Postle 2015, 491) berichtet, dass seit einigen Jahren Neurowissenschaftler die Ansicht vertreten, Emotionen, neurowissenschaftlich betrachtet, seien nicht zu verwechseln mit der subjektiven Erfahrung von Emotion im Individuum, was mit Begriffen einhergeht wie Angst, Eifersucht, Freude, und vielen mehr. Im neurowissenschaftlichen Sinn sollten Emotionen vielmehr als eine Art Überlebenshilfen gedacht werden. Sie sollen dafür sorgen, dass Funktionen und Abläufe wie Verteidigung, Nahrungsaufnahme, Flüssigkeitshaushalt, Thermoregulierung oder Reproduktion reibungslos ablaufen, beziehungsweise zur richtigen Zeit in Gang gebracht werden. Für diese Klassifizierung sind vor allem Arbeiten mit nicht-menschlichen Tieren interessant, die beispielsweise Vergleiche zulassen können, welche Funktionen anderer Säugetiere auch im menschlichen Gehirn zu finden sind.

⁶ Für Literatur zum gegenwärtigen Stand siehe beispielsweise Postle 2015, 309–311.

Die Erforschung von Emotionen erstreckt sich über mehrere Prozesse, von der Aufnahme emotionaler Stimuli über die Kontrolle ihrer Verarbeitung bis hin zu ihrer Auswirkung auf die mentale Gesundheit eines Individuums. Relativ gut bearbeitet ist beispielsweise die Erforschung von Angst. Der experimentelle Psychologe Heinrich Klüver und der Neurochirurg Paul Bucy hatten in den 1930er Jahren Experimente mit nicht-menschlichen Primaten durchgeführt, denen der Temporallappen einschließlich der Amygdala entfernt worden war. Diese und spätere Studien, vor allem seit den 1980er Jahren, legen nahe, dass die Amygdala bei der Verarbeitung von Angststimuli eine große Rolle spielt. Offenbar wurden vor allem Experimente mit Angstkonditionierungen durchgeführt. Ähnlich wie bei den Pavlov'schen Hunden werden hier ein unkonditionierter Stimulus (US) und ein konditionierter Stimulus (CS) miteinander verbunden, bis die Probanden eine Assoziation zwischen beiden aufgebaut haben. Der US ist dabei der ursprüngliche Angstauslöser, beispielsweise ein Elektroschock. Der CS ist ein davon ursprünglich unabhängiger Reiz wie beispielsweise ein bestimmter Ton oder ein Lichtschein. Nach der Konditionierung zeigen die Probanden Angstreaktionen bereits bei der Perzeption des CS allein. Aus Erkenntnissen aus solchen und weiteren Experimenten entstand ein schematisches Diagramm der Arbeit der Amygdala und anderer Hirnregionen bei der Aufnahme und Verarbeitung von CS (Postle 2015, 483). Der eingehende CS wird vom sensorischen Thalamus erfasst und an die Amygdala weitergeleitet. Dort wird er in verschiedenen Zentren verarbeitet und unter anderem an evolutionär ältere Ebenen weitergeleitet, die interne körperliche Reaktionen auslösen: an das zentrale Höhlengrau, das ein Erstarren des Individuums auslöst, sowie an den lateralen Hypothalamus, der das vegetative Nervensystem aktiviert, und an den paraventriculären Nucleus des Hypothalamus, der adäquate Hormone produzieren lässt. Außerdem leitet die Amygdala Reize an das ventrale Striatum weiter, das motorische Reaktionen in Gang setzt. Das sind kontextabhängige Handlungen, die ein Zusammentreffen mit der Bedrohung verhindern, beziehungsweise eine angemessene Verteidigungsreaktion hervorbringen sollen.⁷

Eine auch für den neuroökonomischen Kontext interessante Aufgabe der Amygdala scheint es zu sein, eine schnelle und unbewusste Beurteilung darüber abzugeben, ob ein bestimmtes Gesicht vertrauenswürdig erscheint. In einer Studie (Postle 2015, 478) wurden drei Probanden mit beidseitiger Amygdala-Läsion gebeten, anhand ihnen vorgelegter Fotos ein Urteil darüber abzugeben, ob sie die abgebildeten Personen für vertrauenswürdig hielten. Als Kontrolle

⁷ Dies ist ein sehr grober Überblick über den *Fear Circuit*, der aber für unsere Zwecke ausreichen soll. Für eine detaillierte Beschreibung siehe beispielsweise LeDoux 2012.

dienten die Einschätzungen der selben Gesichter durch eine Gruppe gesunder Probanden. Die drei Amygdala-Patienten neigten dazu, alle vorgelegten Gesichter als vertrauensvoll einzustufen, sogar die 50, die von der Kontrollgruppe als die am wenigsten vertrauenswürdig Erscheinenden eingeschätzt worden waren. Diese Ergebnisse galten aber lediglich für den Durchlauf des Experiments, bei dem den Probanden Fotos vorgelegt worden waren. Bei einem zweiten Durchlauf wurden keine visuellen Stimuli genutzt, die Probanden sollten diesmal anhand ihnen vorgelesener Beschreibungen verschiedener Individuen entscheiden. Hier unterschieden sich die Ergebnisse kaum von denen der Kontrollgruppe. Obwohl die Amygdala nicht zum benachbart gelegenen visuellen System gezählt wird, scheint die Beeinträchtigung der Amygdala-Patienten mit dem visuellen Stimulus zusammenzuhängen. Ralph Adolphs, Daniel Tranel und Antonio Damasio, die Autoren der Studie, kamen daher zu dem Schluss, dass die Amygdala notwendig dafür zu sein scheint, das Hervorholen von Informationen auszulösen, die auf Basis früherer sozialer Erfahrungen bezüglich Klassen von Gesichtern beruhen.

Spätere fMRT-Studien ergaben, dass die Stärke der Amygdala-Aktivität als Antwort auf visuelle Stimulation durch Fotos von Gesichtern mit der Einschätzung der Vertrauenswürdigkeit jedes Gesichtes korrespondiert. Bei einer dieser Studien (Postle 2015, 480) wurde den Probanden gesagt, sie nähmen an einer Studie zum Gesichtsgedächtnis teil. Erst nachdem sie im fMRT-Scanner die Fotos von Gesichtern vorgelegt bekommen hatten, wurden sie nach ihrer Einschätzung zur Vertrauenswürdigkeit der abgebildeten Personen befragt. Ein Vergleich mit der Amygdala-Aktivität während des Scanvorgangs, bei dem die Vertrauenswürdigkeit keine Rolle spielte, zeigte, dass die spätere Einschätzung mit der früheren Aktivität zusammenhing. Diese Erkenntnis wird als Hinweis darauf gedeutet, dass die Amygdala ihre Einschätzung unbemerkt vom Individuum vornimmt. Die Probanden im Scanner formten nicht bewusst ein Bild von der Vertrauenswürdigkeit der fotografierten Person. Die Amygdala scheint ihre spätere Einschätzung aber bereits vorgenommen zu haben. Eine weitere Studie (Postle 2015, 481), die sich mit der Stereotypisierung verschiedener ethnischer Gruppen beschäftigte, kam zudem zu dem Schluss, dass die schnellen, impliziten Einschätzungen der Amygdala gelernt, und nicht festgefahren, sind.

Trotz aller Fortschritte ist die Aufgabe und die Interaktion der Amygdala mit anderen Hirnregionen noch nicht vollständig erforscht. Es gibt noch einige offene Fragen darüber, welche Rolle sie tatsächlich spielt, auch bei der Verarbeitung anderer Stimuli und Emotionen. Generell sind die Emotionen ein weites Forschungsfeld, an dem derzeit weitergearbeitet wird.

Auch an der Erforschung von sozialem Verhalten wird derzeit gearbeitet. Dazu gehört die große Frage nach den neuronalen Grundlagen für eine *Theory*

of Mind (ToM), oder wie wir unser Verhalten an eine Gruppe anpassen, unsere Handlungen kontrollieren oder von anderen Individuen lernen. Auch die Frage nach moralischen Urteilen gehört in dieses Themengebiet. Wie moralische Entscheidungen in unserem Gehirn umgesetzt werden, ist auch interessant für Neuroökonomien, vor allem bei der Untersuchung sozialer Präferenzen. Denn nicht selten ist die Grundlage für ihre Ausprägung ein individuelles Bild von Moralität.

Die Fähigkeit zu moralischen Urteilen tritt gewöhnlich im Alter von 6 bis 7 Jahren hervor, ihre neuronalen Grundlagen sind noch nicht genau erforscht. In einer TMS-Studie (Postle 2015, 462–464) wurde der Zusammenhang von moralischen Urteilen mit dem rechten temporoparietalen Übergang untersucht (dieser war schon früher im Zusammenhang mit *ToM*-Funktionen aufgefallen). In dieser Studie bekamen die Probanden verschiedene Geschichten zu hören, die nach dem gleichen Schema abliefen: Informationen über den Kontext, Andeutungen/Voraussetzungen, Intention und Handlung. Die Intentionen der Protagonisten unterschieden sich in den Geschichten, sie konnten böse oder neutral sein, ebenso wie der Ausgang der Geschichten, den der Empfänger der Handlung zu erleiden hatte, es konnte ein Leid zugefügt werden oder glimpflich ausgehen. Die Probanden sollten sich die Geschichten anhören und dann ein Urteil darüber abgeben, ob die Handlung moralisch erlaubt war oder nicht. Während sie das taten, wurden ihre rechten temporoparietalen Übergänge mit einer TMS-Schleife manipuliert, sodass eine temporäre virtuelle Läsion entstand.⁸ Als Kontrolle diente ein weiterer Durchgang mit der Manipulation einer Kontrollregion, die sich weiter am Hinterkopf befand. Das Ergebnis war, dass unter der Manipulation des temporoparietalen Übergangs die Probanden eine Handlung mit böser Absicht eher als moralisch erlaubt einstufen als bei der Manipulation der Kontrollregion. Dies galt interessanterweise aber nur für die Situationen, in denen eine böse Absicht vorhanden war, aber der betroffenen Person kein Schaden zugefügt wurde. Dies legt für Postle (Postle 2015, 464) den Schluss nahe, dass der rechte temporoparietale Übergang vor allem zur Beurteilung moralisch nicht eindeutiger Situationen beiträgt, in denen eine Diskrepanz zwischen Intention und Ergebnis herrscht.

Allerdings sieht Postle diese Studie kontrovers, da die untersuchte Hirnregion bekanntermaßen auch andere Funktionen erfüllt, die in dieser Studie nicht in Betracht gezogen worden waren. Ebenso gilt wie bei allen TMS-Studien, dass sich die manipulierte Region nicht ganz genau eingrenzen lässt. Umliegende

⁸ Für eine genauere Beschreibung der Funktion von TMS siehe unten, Abschnitt 4.3.4.

Bereiche könnten manipuliert worden sein und so das Ergebnis beeinflusst haben.

Im Allgemeinen ist im Bereich des sozialen Verhaltens zwar einiges Interessantes untersucht worden, verlässliche, allgemein anerkannte Ergebnisse zur Lokalisierung von Emotionsverarbeitungen scheint es aber noch wenige zu geben, sodass auch hier die kognitiven Neurowissenschaften noch weiter bei der Arbeit sind.

4.2.3 Aktuelle Großprojekte in den kognitiven Neurowissenschaften

Unter den zahllosen neurowissenschaftlichen Forschungsprojekten machen vor allem zwei große auf sich aufmerksam. Das eine ist das Human Brain Project, gefördert durch die Horizon 2020-Förderlinie der Europäischen Kommission, das andere die US-basierte BRAIN Initiative, die öffentlich und privat gefördert wird, unter anderem durch die National Institutes of Health und die National Science Foundation. Beginnen wir mit einem Überblick über das Human Brain Project.

Das Human Brain Project ist Europa-basiert und wird von 123 Projektpartnern aus ganz Europa und Israel in einer Vielzahl an kleinen Unterprojekten bearbeitet. Die Projektlaufzeit ist für 2013 bis 2023 angesetzt, die Förderung durch die EU-Kommission liegt bei 1,19 Milliarden Euro. Projektleiter und -initiator ist der bekannte Neurowissenschaftler Henry Markram von der Ecole polytechnique fédérale de Lausanne (EPFL). Er hatte bereits das Vorläuferprojekt Blue Brain Project von 2005 bis 2015 geleitet. Jenes hatte zum Ziel, ein biologisch korrektes, virtuelles Gehirnmodell zu produzieren. Dazu sollten alle Neuronen eines menschlichen Gehirns erfasst und kartiert werden, woraus am Ende eine digitale Simulation des Gehirns entstehen sollte. Als Unterziel dieses monumentalen Vorhabens entstand die Kartierung einer Blue Column eines Ratte Gehirns, also das Modell einer kortikalen Säule einer Ratte von 2 Millimeter Länge und einem Durchmesser von 0,5 Millimeter. Eine solche Säule enthält circa 10.000 Neuronen mit 10^8 Synapsen, eine kortikale Säule gleicher Größe eines Menschen beherbergt circa 60.000 Neuronen. Die Arbeit der Blue Column der Ratte wurde als weiteres Projektziel mit Computern simuliert, wobei ein eigener Rechner pro Neuron verwendet wurde. An diesem Kartierungs- und Simulationsprojekt wird an der EPFL weiterhin gearbeitet, im November 2018 verkündeten die dortigen Forscher die Fertigstellung des ersten digitalen dreidimensionalen

Hirnzellen-Atlas eines gesamten Mäusegehirns.⁹ Das Nachfolgeprojekt Human Brain Project definiert kein so konkretes Gesamtziel, stattdessen heißt es: „The Human Brain Project (HBP) is building a research infrastructure to help advance neuroscience, medicine and computing.“ (humanbrainproject.eu). Diese Weitläufigkeit der teilnehmenden Disziplinen schlägt sich in der Aufteilung des Projektes in sechs Forschungsplattformen nieder (humanbrainproject.eu):

„Neuroinformatics (access to shared brain data), Brain Simulation (replication of brain architecture and activity on computers), High Performance Analytics and Computing (providing the required computing and analytics capabilities), Medical Informatics (access to patient data, identification of disease signatures), Neuromorphic Computing (development of brain-inspired computing) and Neurorobotics (use of robots to test brain simulations).“

Es wird also auf vielen Gebieten gearbeitet. Die Abbildung der Gehirnarchitektur und die Simulation von Gehirnfunktionen sind auch in diesem Projekt ein großes Ziel. Das zweite große Ziel des Projekts ist im Bereich *Neuroinformatics* der Aufbau des Datenrepositoriums Knowledge Graph, in dem die beteiligten Projektgruppen ihre Forschungsdaten untereinander teilen können. Das Repositorium ist bereits in Betrieb, Metadaten sind auch öffentlich zugänglich. Auch im medizinischen Bereich des Projekts steht das Teilen von Daten im Vordergrund: Über die projekteigene Medical Informatics Platform (MIP) sollen Daten wie MRT-Scans von erkrankten und gesunden Gehirnen zwischen Krankenhäusern geteilt werden, um Medizinern Zugang zu breiter Datenlage zu gewähren, die helfen könnte, Patienten vor Ort zu behandeln. Auch von Werkzeugen aus dem Bereich maschinelles Lernen ist die Rede, mit deren Hilfe aus den riesigen Datenmengen neue neurowissenschaftliche Erkenntnisse gewonnen werden sollen. Um unter anderem Bedenken hinsichtlich des Schutzes von Patientendaten zu begegnen, wurde auch ein Ethikkomitee eingerichtet.

Mit der Abbildung und Simulation von Neuronen bis hin zum gesamten Gehirn führt das Human Brain Project die langwierige Arbeit seines Vorgängers weiter. Ein solches Modell des Gehirns zu besitzen, wäre sicherlich förderlich für die kognitiven Neurowissenschaften, auch bereits wegen der Erkenntnisse, die auf dem Weg dorthin gewonnen werden dürften. Schon allein aufgrund der Vielzahl an Neuronen im menschlichen Gehirn dürfte es jedoch noch eine Weile dauern, bis ein Modell fertiggestellt ist, noch länger bis die Funktionen verstanden und simuliert werden können. Vielleicht ist dieses also nicht das letzte Projekt, das sich mit dieser Aufgabe beschäftigen wird. Mit dem Aufbau von

⁹ Siehe <https://bluebrain.epfl.ch/op/edit/page-158777.html>, abgerufen am 22.01.2019

Plattformen zum Teilen von Forschungs- und Patientendaten liegt das Projekt jedenfalls genau im aktuellen Trend der empirischen Wissenschaften. Allerdings werden noch einige vor allem rechtliche Hürden überwunden werden müssen, bis eine solche Plattform betrieben werden könnte, besonders in internationaler Zusammenarbeit.

Das zweite große Projekt, das wir hier betrachten, ist die BRAIN Initiative. Die Abkürzung steht für Brain Research through Advancing Innovative Neurotechnologies. Das Projekt ist US-basiert und Teil der breiteren White House Neuroscience Initiative. Die BRAIN Initiative hat eine Laufzeit von 2013 bis 2025 und wird sowohl öffentlich als auch privat finanziert. Die privaten Förderer sind das Allen Institute for Brain Science, das Howard Hughes Medical Institute und The Kavli Foundation. Öffentliche Gelder kommen von den National Institutes of Health (NIH), der National Science Foundation (NSF) und bemerkenswerterweise von der Defence Advanced Research Projects Agency (DARPA), einer Behörde des Verteidigungsministeriums, die Forschungen für das Militär finanziert. Insgesamt erhält die BRAIN Initiative einen Etat von 4,9 Milliarden US-Dollar, mit einer jährlichen Basisfinanzierung von circa 300 Millionen US-Dollar. Die Ziele der Initiative klingen nicht so weit von denen des Human Brain Project entfernt (braininitiative.nih.gov):

„The Brain research through Advancing Innovative Neurotechnologies® (BRAIN) Initiative is aimed at revolutionizing our understanding of the human brain. By accelerating the development and application of innovative technologies, researchers will be able to produce a revolutionary new dynamic picture of the brain that, for the first time, shows how individual cells and complex neural circuits interact in both time and space. Long desired by researchers seeking new ways to treat, cure, and even prevent brain disorders, this picture will fill major gaps in our current knowledge and provide unprecedented opportunities for exploring exactly how the brain enables the human body to record, process, utilize, store, and retrieve vast quantities of information, all at the speed of thought.“

Auch hier werden die medizinischen Aspekte und Vorteile der Forschungen in den Vordergrund gerückt. Auch hier soll ein Modell des Gehirns und seiner Funktionsweise entstehen. Es wird aber auch ein großer Fokus auf technologische Entwicklungen gelegt, abseits des Aufbaus von Repositorien. Die ersten fünf Jahre der Initiative sollen vor allem technologische Entwicklungen fördern, wie beispielsweise neue Diagnoseverfahren oder neue bildgebende Instrumente. Wie wir in Abschnitt 4.3 sehen werden, haben die einzelnen neurowissenschaftlichen Technologien wie funktionale Magnetresonanztomographie (fMRT) oder die Elektroenzephalographie (EEG) methodische und technische Nachteile. Die

BRAIN Initiative unterstützt daher Forschungsprojekte, die sich mit der Suche nach neuen Möglichkeiten beschäftigen, das Gehirn und seine Abläufe sichtbar beziehungsweise messbar zu machen. Nach den ersten fünf Jahren sollen dann vor allem Projekte gefördert werden, die auf neurowissenschaftlichen Erkenntnisgewinn ausgerichtet sind. Insgesamt wird gefördert in den Bereichen *Data Coordination/Informatics, Human Imaging, Neural Recording, Cell/Circuit Tools, Cell Types, Understanding Circuits, Training* und *Neuroethics*. Die zehn Institute der NIH sind Projektpartner, um Förderung bewerben können sich aber offenbar Forschungsvorhaben jeder amerikanischen Forschungseinrichtung.

Mit einem Budget, das noch weit über der Rekordsumme der EU-Kommission für das Human Brain Project liegt, können zahlreiche Projekte finanziert werden. Auch hier ist die Arbeit sicher sehr klein- und vierteilig, bis ein Modell des Gehirns entsteht. Es ist sicherlich auch eine gute Idee, die Suche nach neuen Diagnose- und Messverfahren zu unterstützen. Bedenkt man, welche Forschungs- und Diagnosemöglichkeiten sich den Neurowissenschaften durch die Entwicklung bildgebender Verfahren in den 1990er Jahren eröffneten, könnten weitere Instrumente und Techniken den Erkenntnisgewinn in den Kognitionswissenschaften weiter vorantreiben. Bedenklich allerdings ist die Finanzierung durch die DARPA. Man will sich gar nicht ausmalen, welche Möglichkeiten ein Modell des Gehirns und seiner Funktionalitäten sowie neue Technologien, es zu manipulieren, sich für das Militär eröffnen werden. Betrachten wir im folgenden Unterkapitel aber zunächst eine Auswahl an derzeit zur Verfügung stehenden Techniken und Instrumenten mit ihren Funktionsweisen sowie Vor- und Nachteilen.

4.3 Empirische Methoden

Für ihre Untersuchungen und Experimente stehen den kognitiven Neurowissenschaftlern eine ganze Reihe an verschiedenen Verfahren zur Verfügung. Diese unterscheiden sich nicht nur in ihren Techniken und Anwendungsgebieten, sondern auch in Komplexität und Kosten. Jedes besitzt Vor- und Nachteile und eignet sich für eine bestimmte Aufgabe besonders gut oder schlecht. Daher werden verschiedene Verfahren im besten Fall ergänzend eingesetzt, sodass in mehreren Studien unterschiedliche Blickwinkel auf bestimmte Fragen ermöglicht werden können.

Im Folgenden werfen wir einen Blick auf ausgewählte Verfahren, die in den kognitiven Neurowissenschaften regelmäßig angewandt werden und auch in neuroökonomischen Experimenten Verwendung finden. Hierbei werden invasive von

nicht-invasiven Methoden unterschieden. Bei Letzteren wird der Schädel geöffnet, um direkte Messungen zu erlauben. Diese Art von Experimenten wird ausschließlich bei Versuchstieren angewandt. Für die Untersuchung menschlicher Versuchspersonen wurden in den vergangenen Jahrzehnten Messmethoden entwickelt, die neuronale Aktivitäten von außerhalb des Schädels ermitteln. Sie werden teilweise auch für Versuche an nicht-menschlichen Primaten verwendet, sind allerdings in ihrer Messung nicht so präzise wie invasive Methoden. In dieser Aufstellung betrachten wir zunächst die in der heutigen Forschung geläufigen nicht-invasiven Methoden mit ihren jeweiligen Anwendungsgebieten und Vor- und Nachteilen, bevor wir zur invasiven Methode der direkten Abnahme von Aktionspotenzialen und ihrer Anwendung in Tierexperimenten kommen.

4.3.1 Funktionelle Magnetresonanztomographie (fMRT)

Mit Verfahren der funktionellen Bildgebung wie die funktionelle Magnetresonanztomographie (fMRT) oder die Positronenemissionstomographie (PET, s. u.) können neurale Aktivitäten im zentralen Nervensystem gemessen werden, ohne den Schädel zu öffnen. Vor dem Hintergrund der schlechten Datenlage im 19. und 20. Jahrhundert war die Entwicklung funktioneller Bildgebungsverfahren in den 1990er Jahren eine große Bereicherung für Humanmedizin und neurowissenschaftliche Forschung.

Sowohl fMRT als auch PET messen neuronale Aktivitäten indirekt über Veränderungen im Stoffwechsel. Wo Nervenzellen aktiviert sind, ändert sich der Stoffwechsel in ihrer Umgebung, was den regionalen zerebralen Blutfluss ansteigen lässt. Dieser Anstieg wird gemessen und so ein Bild davon generiert, welche Nervenzellverbände wann im Verlauf eines Versuchsdurchlaufs aktiviert sind. (Büchel et al. 2012, 10). Aufgrund dieser Arbeitsweise können funktionelle Bildgebungsverfahren keine klare Auskunft darüber geben, ob eine Hirnregion eine bestimmte Aktion begründet oder ob sie allein einen kognitiven Prozess steuert. Sie können aber aufzeigen, dass Aktivität in bestimmten Hirnregionen oft mit bestimmten Stimuli einhergeht oder dass Information in bestimmten Hirnregionen repräsentiert ist, bewusst oder unbewusst, und können Anzeichen dafür geben, wo beschädigte Gehirne nicht normal arbeiten. (Carter et al. 2010, 13–14).

Um die Stoffwechselveränderungen in der Nähe aktivierter Neuronen zu messen, macht sich die fMRT den *Blood Oxygen Level-Dependent Effect* (BOLD-Effekt) zunutze. Der BOLD-Effekt beschreibt die Veränderung der Sauerstoffkonzentration im Blut bei Aktivität von Nervenzellen: Werden Nervenzellen aktiviert, erhöht sich ihr Zellstoffwechsel und sie benötigen mehr Sauerstoff, den sie aus

den umliegenden Blutgefäßen beziehen. Durch den erhöhten Verbrauch steigt in den Blutgefäßen die Konzentration von deoxygeniertem Hämoglobin, während die Konzentration an oxygeniertem Hämoglobin sinkt. In Sekunden reagiert die Blutversorgung des Gehirns mit der Erhöhung des regionalen zerebralen Blutflusses, was sauerstoffreiches Blut zu der betroffenen Hirnregion bringt. Die Blutgefäße um aktivierte Neuronen herum weisen also eine höhere Konzentration an oxygeniertem Hämoglobin auf als Blutgefäße um Neuronen im Ruhezustand. Gleichzeitig sinkt auch die Konzentration an deoxygeniertem Hämoglobin wieder. Da deoxygeniertes Hämoglobin paramagnetische Eigenschaften besitzt, ändert sich in der betroffenen Hirnregion die Reaktion des Blutes auf Magnetfelder. (Carter et al. 2010, 14; Büchel et al. 2012, 12). Kurz gesagt ist sauerstoffreiches Blut weniger stark magnetisch als sauerstoffarmes. Diese Änderungen werden vom Magnetfeld des fMRT-Scanners erfasst und farblich kontrastiert dargestellt.

Die Technik eines fMRT-Scanners beruht auf der Technologie von Magnetresonanztomographen (MRT), wie sie heute unter anderem in Krankenhäusern weit verbreitet sind. Der Scanner besteht aus einer langen Röhre, in der magnetische Spulen untergebracht sind. Patienten oder Versuchspersonen werden auf einer Liege in die Röhre geschoben. Die MRT-Technik macht sich zunutze, dass sich die Wasserstoffprotonen in den Körpern der Patienten um die eigene Achse drehen und dadurch ein winziges magnetisches Feld entsteht. Die Magnetfelder der einzelnen Protonen sind normalerweise zufällig ausgerichtet und heben sich größtenteils gegenseitig auf. Der Scanner erzeugt nun ein Magnetfeld längs zum Körper des Patienten, von den Füßen in Richtung des Kopfes. Dieses Magnetfeld richtet die Wasserstoffprotonen aus, manche parallel zum Magnetfeld (Richtung Kopf), manche antiparallel (Richtung Füße). Die Mehrheit der Protonen wird parallel ausgerichtet, da dies energetisch leicht vorteilhaft für sie ist. Durch dieses Ungleichgewicht entsteht im Körper des Patienten ein netto magnetischer Feldvektor (*Net Magnetisation Vector NMV*) in paralleler Richtung zum Magnetfeld des Scanners. Die Protonen bleiben dabei nicht stationär im Körper, sondern präzedieren um ihre eigene Achse, ähnlich wie ein Kreisel. Die Kreiselfrequenz hängt dabei von der Stärke des Magnetfeldes ab, in dem sie sich befinden. Je stärker das Magnetfeld, desto höher die Frequenz. Gewöhnliche MRT-Scanner erzeugen Feldstärken von 1,5 bis 3,0 Tesla, wobei in jüngster Zeit auch Geräte mit 7,0 oder gar 9,4 Tesla eingesetzt werden. Höhere Feldstärken ergeben eine bessere räumliche Auflösung und erhöhte Kontraste. Gleichzeitig steigt aber auch die Frequenz, mit der die Atome des Patienten zum Schwingen angeregt werden, sodass die Energiedichte, die der Körper des Patienten aufnehmen muss, ansteigt. Das kann zu Unwohlsein bis hin zu lokalen Verbrennungen führen. (Carter et al. 2010, 6–7).

Um nun ein Bild zu erzeugen, werden zusätzlich zum längs anliegenden Magnetfeld Hochfrequenzpulse (*RF Pulses*) in der gleichen Frequenz wie die Kreiselfrequenz der Protonen auf den Patienten gerichtet. Dadurch werden zwei Prozesse in Gang gebracht: Erstens nehmen einige Protonen in paralleler Ausrichtung die Energie auf, kehren ihre Polarität um zu antiparallel, und verringern so die Nettomagnetisierung im Körper. Zweitens regt der RF-Puls einige Protonen zu einer Rotation an, die sich zu einem Vektor quer zum externen Magnetfeld (also senkrecht zum Körper des Patienten) summiert. Somit wird eine neue, quer anliegende Magnetisierung aufgebaut. Nachdem der RF-Puls ausgeschaltet wurde, verlassen die hochenergetisierten Protonen ihre Kreiselfrequenz und kehren zu ihrer vorherigen Längsausrichtung zurück. Dadurch steigt die Längsmagnetisierung im Körper wieder auf ihren ursprünglichen Wert zurück, während gleichzeitig die Quermagnetisierung auf Null absinkt. Beide Vorgänge werden erfasst: Die Zeit, in Millisekunden, die ein bestimmter Prozentsatz der Protonen benötigt, um zur Längsausrichtung zurückzukehren, wird T1 genannt. Die Relaxationszeit der Quermagnetisierung wird T2 genannt. Diese beiden Werte sind wichtig für die Bildgebung, da unterschiedliche Gewebesubstanzen unterschiedliche Werte für T1 und T2 relativ zueinander aufweisen. So ist das T1-Signal von weißer Substanz stärker als das von grauer Substanz, während das T1-Signal von grauer Substanz stärker ist als das von Zerebrospinalflüssigkeit. Beim T2-Signal verhält es sich für diese Substanzen genau andersherum. Der Computer des MRT-Gerätes berechnet ein Bild des Gehirns anhand dieser unterschiedlichen Signale. Solche schwarz-weißen Bilder sind auch von diagnostischen Untersuchungen in Krankenhäusern bekannt. (Carter et al. 2010, 9–10).

Das fMRT nutzt die MRT-Technologie plus den BOLD-Effekt, um nicht nur statische Bilder der Materie zu produzieren, sondern auch die Aktivierungen der Nervenverbände anzuzeigen. Diese Aktivierungen werden dann zur Veranschaulichung bunt eingefärbt, mit verschiedener Farbgebung für verschiedene Aktivierungsgrade.

Einer der Vorteile von fMRT-Geräten ist sicherlich, dass sie das gesamte Gehirn in verschiedenen wählbaren Schichten erfassen können. Andere Verfahren, wie beispielsweise Elektroenzephalographie (EEG, s. u.) können nur Aktivierungen an der Oberfläche des Hirngewebes messen. Gleichzeitig ist die Methode nicht invasiv und kommt ohne Injektion radioaktiver Substanzen aus, und kann daher beliebig oft angewandt werden. Damit ermöglicht die fMRT die Erforschung der neuralen Grundlagen für Kognition, Emotionen und sensorische Empfindungen an lebenden Menschen wie keine andere Technik (Carter et al. 2010, 14). Zudem liefert die auf dem BOLD-Effekt basierende Technik eine

räumliche Auflösung im Bereich von Millimetern, was sehr gut ist (Büchel et al. 2012, 12).

Allerdings hat der BOLD-Effekt auch den Nachteil, dass er zeitlich verspätet nach der eigentlichen Aktivierung der Nervenzellen auftritt, ca. 5 Sekunden (Büchel et al. 2012, 11). Das bedeutet, dass Messungen auf Basis des BOLD-Effekts eine relativ schlechte zeitliche Auflösung besitzen. Hier sind elektrophysiologische Verfahren wie die EEG im Vorteil, wie wir in Abschnitt 4.3.3 sehen werden.

Außerdem muss bei der fMRT-Technik auch bedacht werden, dass die Verlängerung der T2-Relaxationszeit durch den BOLD-Effekt das MR-Signal bei einem gewöhnlichen 1,5 Tesla-Scanner nur um wenige Prozent erhöht. Die Gefahr, eine Aktivierung zu übersehen, ist also hoch, ebenso wie die Gefahr, gemessene Aktivierungen fälschlicherweise mit dem Versuchsstimulus in Verbindung zu bringen. Die Analysten können nicht sicher sein, dass der Scanner alle relevanten Aktivierungen erfasst hat. Genauso wenig können sie sicher sein, dass die erfassten Aktivierungen tatsächlich mit dem Versuchsstimulus zusammenhängen. Denn schon kleine Veränderungen im Sauerstoffgehalt des Blutes können vom Scanner als Aktivierung erkannt werden. Bei der Auswertung der Scannerdaten spricht man daher von einem Rauschen, das die ständigen (Hintergrund-)Aktivitäten des Gehirns widerspiegelt. Denn während eines Versuchs sind nicht nur Nervenzellverbände aktiv, die auf die Versuchsstimuli reagieren, sondern auch solche Zellen, die auf viele andere Reize reagieren, die nicht direkt mit der Experimentaufgabe zusammenhängen. Das Gehirn ist ständig aktiv, empfängt und verarbeitet ständig Informationen und gibt ständig Anweisungen aus. Eine wichtige Aufgabe für die Forscher ist es daher, die durch den Versuch verursachten Aktivierungen aus dem Rauschen herauszufiltern. Dies wird über spezielle Experimentdesigns zu erreichen versucht. Beispielsweise werden sogenannte Kontraste in den Versuch eingebaut. Die Probanden bekommen zusätzlich zur Versuchsaufgabe weitere Aufgaben, die in bestimmter Weise von den Versuchsaufgaben verschieden sind. Die Aktivierungsmuster von Versuchsaufgabe und Kontrast-Aufgabe werden dann subtrahiert, um die Aktivierungen zu erhalten, die allein durch die Versuchsaufgabe bedingt wurden. Inzwischen hat sich eine ganze Reihe an verschiedenen Möglichkeiten und Versuchsanordnungen in Lehrbüchern bewährt, mit denen wir uns in Abschnitt 5.2 an einem Beispielperiment näher befassen werden.¹⁰

Zusätzlich werden neurowissenschaftliche und neuroökonomische Experimente mit fMRT mit jeder Person zweimal durchgeführt. Durch die doppelte

¹⁰ Für eine Übersicht über verbreitete Experimentdesigns siehe Karnath et al. (2012) oder Hermey et al. (2010).

Datenlage soll die Verlässlichkeit der Analyse verbessert werden. Der Nachteil dabei ist, dass die Gefahr besteht, dass die Probanden bei einem zweiten Durchlauf aus dem ersten Durchgang gelernt haben und nun anders reagieren könnten. Dieser Lernaspekt kann die Ergebnisse des zweiten Durchgangs verfälschen und damit auch die Ergebnisse und Schlussfolgerungen der gesamten Studie.

Ein weiterer Nachteil der fMRT sind die hohen Kosten. Gerätepreise liegen im einstelligen Millionenbereich, je nach Stärke der Magneten. Hinzu kommen die Kosten für jeden Scan, weshalb nicht alle Krankenhäuser oder neurowissenschaftlichen Forschungseinrichtungen im Besitz von fMRT-Geräten sind. Daher ist es nicht unüblich, dass Neurowissenschaftler für ihre Studien Scanner-Betriebszeiten an einem Krankenhaus mieten. Nicht selten finden fMRT-Studien dort nachts statt, da zu diesen Zeiten die Belegung des Scanners durch das Krankenhaus selbst am niedrigsten ist. Für jede Studie muss abgewogen werden, wie viele Probanden sie umfassen soll. Mehr Probanden bedeuten mehr Daten, die die Analysen und Schlussfolgerungen auf eine breitere Basis stellen würden. Mehr Probanden bedeuten aber auch mehr Scanner-Sitzungen, zumal jede Person zweimal gescannt werden muss, und somit auch mehr Kosten. Nicht zuletzt aus diesem Grund umfassen fMRT-Studien häufig weniger Probanden als Studien mit anderen Messtechniken.

4.3.2 Positronenemissionstomographie (PET)

Wie wir gerade gesehen haben, ermöglicht fMRT die Untersuchung von Nervenaktivitäten im menschlichen Gehirn und dient daher der Erforschung der neuronalen Basis von Kognition. Was es aber nicht aufzeigen kann, ist die neurochemischen Zusammensetzung der neuronalen Aktivitäten. Um beispielsweise zu untersuchen, welche Neurotransmitter Veränderungen in den Nervenzellaktivitäten herbeiführen, wird unter anderem die Positronenemissionstomographie (PET) verwendet.

Auch die PET misst Aktivitäten indirekt. Je nachdem, welche chemischen Substanzen untersucht werden sollen, gibt es verschiedene Vorgehensweisen. Generell werden bei einem PET-Experiment oder einer -Untersuchung bestimmte Isotope beispielsweise in die Halsschlagader der Patienten injiziert. Sie verbreiten sich daraufhin im Gehirn des Patienten. Wenn sie zerfallen, wird ein Positron freigesetzt, das schnell auf ein Elektron stößt. Bei ihrem Zerfall werden zwei Gammaphotonen ausgesendet, die sich in einem Winkel von 180° voneinander entfernen. (Büchel et al. 2012, 10). Innerhalb von Nanosekunden verlassen sie den Körper und werden von Gammastrahlendetektoren erfasst, die in einer Röhre

ringförmig um den Kopf des Patienten angebracht sind. Durch die Erfassung zweier, sich im 180°-Winkel voneinander weg bewegender Photonen zur gleichen Zeit, berechnet ein Computer den Ursprungsort des Zerfalls im Gehirn. Durch viele solcher Messungen entsteht ein Bild vom Verteilungsmuster der Zerfälle und dadurch der Verteilung des injizierten Isotops im Gehirn.

Um verschiedene neurochemische Prozesse zu untersuchen, werden verschiedene Isotope eingesetzt. So können mit Isotopen markierte Liganden an spezifische Rezeptoren binden und ihre Aktivität so für den Detektor sichtbar machen. Carter und Shieh (Carter et al. 2010, 19) erklären:

„For example, a radioactive ligand that binds to serotonin receptors can indicate the locations and binding potential of these receptors in the brain, providing information about the relative metabolism of serotonin in human subjects.“

Die Autoren betonen, dass mit der PET die Verstoffwechselung spezifischer bioaktiver Moleküle untersucht werden kann. Damit bietet die PET Einsichten, die eine fMRT-Untersuchung nicht bieten kann. (Carter et al. 2010, 19). Auch der regionale zerebrale Blutfluss kann mit einer PET gemessen werden. Dazu wird häufig radioaktiv markiertes Wasser (H_2^{15}O) in die Halsschlagader injiziert. Durch die kurze Halbwertszeit von circa 2 Minuten kann der Blutfluss zu aktivierten Nervenzellen nachvollzogen werden. (Büchel et al. 2012, 10). Außerdem kann die Aktivierung von Nervenzellen über ihren erhöhten Stoffwechsel gemessen werden. Dazu wird meist Fluorodeoxyglucose (FDG) verwendet, eine radioaktive Form von Glucose. Da aktivierte Nervenzellen mehr Energie benötigen, also Glucose, kann eine erhöhte Aufnahme von Glucose aus dem Blut ein Indikator für aktive Nervenzellen sein. (Carter et al. 2010, 18).

Die PET bietet Einsichten in die neurochemischen Prozesse menschlicher Gehirne. Ähnlich wie die fMRT ist auch dies eine wertvolle Erkenntnisquelle für kognitive Neurowissenschaftler, da sie nicht invasiv ist und an lebenden menschlichen Probanden angewandt werden kann. Dabei bietet sie eine ähnliche zeitliche Auflösung wie fMRT.

Allerdings ist der Einsatz von PET auch sehr teuer, teurer noch als der Einsatz von fMRT. Das hängt nicht nur mit dem Gerät, dem Detektor und dem Computer zusammen, sondern auch mit den radioaktiven Substanzen, die benötigt werden. Da sie eine kurze Halbwertszeit haben müssen, können sie nicht außerhalb der Klinik oder Forschungseinrichtung hergestellt werden. Während des Transports zum PET würden sie bereits zerfallen. Daher müssen die Isotope in einem eigenen Zyklotron vor Ort hergestellt werden. Die Anschaffung und Instandhaltung dieser Einrichtung verursacht nicht unerhebliche Zusatzkosten. (Carter et al. 2010, 19).

Hinzu kommt die nötige Injektion radioaktiver Substanzen in die Körper der Patienten. Manchmal sogar mehrere gleichzeitig, je nach Untersuchungsziel. Die Strahlenbelastung ist nicht nur ein Gesundheitsrisiko, sie begrenzt auch die möglichen Wiederholungen der Experimente oder Untersuchungen an einem Patienten. Die fMRT dagegen benötigt keine radioaktiven Substanzen und ist theoretisch beliebig oft an einem Patienten durchführbar. Da sie eine ähnliche Auflösung bietet, wird für kognitive Forschungen für gewöhnlich die fMRT bevorzugt. Der hohe Aufwand und die Belastung für die Patienten lohnen einen Einsatz der PET in solchen Fällen nicht. (Büchel et al. 2012, 10).

4.3.3 Elektroenzephalographie (EEG)

Die Elektroenzephalographie (EEG) zählt zu den elektrophysiologischen Verfahren der kognitiven Neurowissenschaften. Im Gegensatz zu den bildgebenden funktionellen Verfahren fMRT und PET messen sie neuronale Aktivitäten nicht indirekt über Stoffwechselprozesse, sondern erfassen postsynaptische Membranpotenziale oder Aktionspotenziale mehrerer aktivierter Nervenzellen. Dazu werden Elektroden an der Kopfhaut befestigt, die die Summe der Potenziale im Gehirn erfassen. Häufig werden dafür den Probanden Hauben oder Netze aufgesetzt, auf denen die Elektroden angebracht sind. Um die Übergangswiderstände zu verringern und die Messungen zu verbessern, wird vorher ein leitfähiges Gel auf die Kopfhaut aufgetragen. Die Elektroden messen dann die Potenziale differenziell mit Hilfe einer zusätzlichen Referenzelektrode. Sie muss daher so angebracht werden, dass sie selbst keine Gehirnströme messen kann, meist um die Nasen der Probanden. Da die Spannungen von Aktions- und postsynaptischen Potenzialen im Mikrovoltbereich liegen, werden die Messdaten von einem Verstärker aufbereitet. (Büchel et al. 2012, 24–25).

Anders als fMRT und PET liefert EEG keine bildliche Darstellung des Gehirns, kann aber bestimmte Zustände im Gehirn anzeigen. Carter und Shieh (Carter et al. 2010, 20) vergleichen die Leistungen der EEG mit einem Mikrofon, das über einer großen Menschenmenge angebracht wird, wie beispielsweise in der Silvesternacht über dem Times Square in New York. Das Mikrofon nimmt die Geräusche der Menge auf. Dabei ist nicht auszumachen, welches Individuum welche Geräusche macht, ebenso kann das Mikrofon nicht nur auf ein Individuum gerichtet werden. Jedoch ist es mit dieser Methode möglich, auszumachen, wann etwas Wichtiges geschieht, in unserem Beispiel die Jubel der Menschen um Mitternacht. Analog können mit einer EEG-Messung

nicht die Aktivitäten einzelner Neuronen oder kleinerer Neuronenverbände separat gemessen werden. Aber es ist möglich, festzustellen, wann ein bedeutsames Ereignis eintritt, wie beispielsweise wenn der Proband in einem Experiment einen Stimulus wahrnimmt.

Dies mag nicht wie eine präzise Darstellung der Ereignisse in einem Gehirn erscheinen, dennoch kann eine EEG wertvolle Informationen liefern. Da sie Neuronenaktivitäten nicht indirekt über Stoffwechselprozesse misst, die erst *nach* einer Aktivität auftreten, sondern direkt über elektrische Potenziale, die *während* einer Aktivität auftreten, liefert die EEG eine sehr gute zeitliche Auflösung im Bereich von Millisekunden. Vor allem in Kombination mit der sehr guten räumlichen Auflösung aus fMRT- oder PET-Messungen, kann die EEG präzise Daten über die Arbeit des Gehirns liefern.

4.3.4 Transkranielle Magnetstimulation (TMS)

Für eine transkranielle Magnetstimulation (TMS) wird elektrischer Strom mit sehr hoher Stromstärke in eine portable Spule eingeleitet, wodurch in der Spule ein starkes Magnetfeld entsteht. Die Spule wird dann nahe an den Schädel des Probanden gebracht. Durch das starke Magnetfeld entsteht an der Oberfläche des Gehirns ein schwacher elektrischer Strom, der Neuronenverbände auf der Kortexoberfläche reizt. Auf diese Weise können Gehirnareale gezielt stimuliert werden, beispielsweise der motorische Kortex, was zum Beispiel ein Zucken der Finger auslöst. Die stimulierende Wirkung der TMS kann auch inhibitorisch genutzt werden, indem sie wiederholt auf die selbe Stelle angewandt wird. Dadurch wird diese Region selektiv gestört, wodurch eine virtuelle funktionelle Läsion entsteht. (Büchel et al. 2012, 27). Das ist beispielsweise dann von Vorteil, wenn keine oder nicht genügend Patienten mit natürlichen Läsionen in den zu untersuchenden Hirnregionen zur Verfügung stehen.

Die TMS gehört zu den Stimulationsverfahren und wird häufig als Läsionsmodell verwendet. Läsionsmodelle sind ein wichtiger Bestandteil der Erforschung der Aktivierungen und des Zusammenspiels verschiedener Hirnregionen bei der Bearbeitung kognitiver Aufgaben. Wird beispielsweise bei einer fMRT-Studie eine bestimmte Hirnregion als für die Experimentaufgabe funktionell relevant angesehen, weil sie in den entscheidenden Momenten aktiv war, so sind die beobachteten Aktivitäten allein kein Beweis für die tatsächliche Relevanz dieser Hirnregion. Es bedarf weiterer Messungen auf anderer Grundlage, um die Relevanz zu bestätigen. (Büchel et al. 2012, 26). Läsionsmodelle gehen dabei

sozusagen von einer anderen Seite an die Frage heran, ob bestimmte Hirnregionen für bestimmte Aufgaben relevant sind. Sie zeigen nicht an, welche Hirnareale bei welcher Aufgabe aktiv sind, sondern helfen zu erforschen, welche Hirnareale für diese Aufgabe unverzichtbar sind (Büchel et al. 2012, 16). Für solche Studien werden Patienten gesucht, die nach einem Unfall oder einer Erkrankung, wie beispielsweise einem Schlaganfall, eine bestimmte Einschränkung ihrer kognitiven Leistungsfähigkeit aufweisen. Lange Zeit war die Untersuchung natürlicher Läsionen die einzige Bezugsquelle für Probanden. Mit Hilfe der TMS können auch gesunde Probanden im Rahmen von Läsionsmodellen untersucht werden.

Ein großer Nachteil der TMS ist allerdings, dass sie, wie auch die EEG, auf die oberflächennahen Nervenverbände beschränkt ist. Zudem ist die räumliche Auflösung der TMS nicht sehr hoch, da das Magnetfeld räumlich ausgedehnt ist. Dadurch ist es schwierig, nur ganz bestimmte Gehirnareale zu stimulieren beziehungsweise zu stören. Auch benachbarte Areale werden leicht von dem Magnetfeld erfasst.

4.3.5 Ableitung von Aktionspotenzialen

Die bisher genannten Methoden der kognitiven neurowissenschaftlichen Forschung haben den Vorteil, dass sie nicht invasiv sind und an lebenden, gesunden Menschen angewandt werden können. Dieser Vorteil hat allerdings den Preis geringerer räumlicher (EEG, TMS) oder zeitlicher Auflösung (fMRT, PET). Die Ableitung von Aktionspotenzialen dagegen liefert eine größtmögliche räumliche und zeitliche Auflösung. Sie ist eine invasive Methode und wird in der kognitiven neurowissenschaftlichen Forschung ausschließlich an Tieren angewandt.

Tierexperimente werden in den kognitiven Neurowissenschaften zum größten Teil mit Nagetieren oder nicht-menschlichen Primaten durchgeführt. Der Vorteil von Affen als Probanden ist ihre phylogenetische Nähe zum Menschen, die einen ähnlichen Aufbau der Gehirne bedeutet. So lassen sich Forschungsergebnisse aus Studien mit beispielsweise Schimpansen leichter auf den Menschen übertragen. Die Vorteile von Nagetieren als Probanden sind ihre einfachere Haltung und geringere Generationszeit, was vor allem für Studien mit Züchtungen von Bedeutung ist. (Büchel et al. 2012, 28). Da Nagetiere ebenso wie Menschen Säugetiere sind, sind ihre organischen Übereinstimmungen groß genug, um für bestimmte kognitive Leistungen Rückschlüsse auf menschliche Gehirne zuzulassen.

Auch bei der in Abschnitt 4.3.3 betrachteten Elektroenzephalographie werden Aktionspotenziale im Gehirn gemessen, allerdings durch die Barriere der Schädelknochen hindurch. Bei der invasiven Ableitung von Aktionspotenzialen

werden diese direkt an den Neuronen abgenommen. Hierzu wird den Versuchstieren unter Narkose eine kleine Öffnung in den Schädel gebohrt, in die eine Ableitkammer implantiert wird. Das ist für gewöhnlich ein kleiner Kubus aus Kunststoff, der mit einem Deckel verschlossen und geöffnet werden kann. Auf diese Weise ist ein Zugang zum Gehirn des Tieres jederzeit möglich und Reinigung und Desinfektion der Öffnung werden erleichtert. Um die Aktionspotenziale im Gehirn zu messen, werden dünne Elektroden in das Gehirn eingeführt, entweder vor jeder Experimentsitzung neu und danach wieder entfernt, oder sie werden chronisch implantiert. (Büchel et al. 2012, 29). Für die Messungen gibt es verschiedene Möglichkeiten: Für *Single-Unit*-Messungen werden drahtfeine Elektroden in die erforschte Gehirnregion eingeführt. Sie messen die Aktionspotenziale mehrerer Neuronen in der nächsten Umgebung, die zu einer einzelnen funktionalen Einheit gehören (Ruff et al. 2014, 80). Die zweite Möglichkeit ist die Verwendung der Patch-Clamp-Technik, bei der für die Messungen Mikropipetten verwendet werden, deren Öffnungen einen Durchmesser von 1 μm und weniger aufweisen. Durch einen leichten Unterdruck in der Pipette werden die Spitzen mit der Zellmembran eines Neurons verbunden, was auch das Umgebungsrauschen, verursacht durch umliegende Neuronen, weitgehend abschirmt. Diese Methode erlaubt es auch, Systeme aus mehreren hundert Mikroelektroden einzusetzen, die viele Messungen gleichzeitig liefern. (Büchel et al. 2012, 29).

Die derart ausgestatteten Tiere werden dann mit unterschiedlichen Experimentaufgaben und -stimuli konfrontiert, wobei ihre Reaktionen in Aktionspotenzialen gemessen werden. Die Experimente variieren stark in Aufbau und Aufgaben, je nach Art der Tiere. Während die Untersuchung von Nagetieren für basale kognitive Aufgaben interessant ist, können mit bestimmten nicht-menschlichen Primaten komplexe kognitive Fertigkeiten untersucht werden, die bekannten Spielen aus ökonomischen Laboren ähneln. Dafür werden die Tiere monatelang täglich darauf trainiert, bestimmte Aufgaben zu lösen und verlässlich mitzuwirken. Denn die Forscher sind bei solchen Experimenten auf die Mitarbeit der Tiere angewiesen, ihre Kooperation kann nicht erzwungen werden. Daher werden die Tiere mit Belohnungen, meist Fruchtsäfte oder Obst, zur Mitarbeit angeregt. Sind sie gesättigt oder unmotiviert, werden die Experimente unterbrochen. Trainierte, kooperationsbereite Tiere sind sehr wertvoll und werden gut gepflegt, damit sie jahrelang an den Experimenten teilnehmen können.

Daran lässt sich bereits ein großer Nachteil von invasiven Ableitungen von Aktionspotenzialen und Tierexperimenten im Allgemeinen erahnen: Die artgerechte Haltung, die Pflege und das Training der Tiere ist langwierig und kostenintensiv, vor allem bei nicht-menschlichen Primaten. Zusätzlich ist die Präzision der Abnahme von Aktionspotenzialen ihr größter Vorteil, aber auch ein

großer Nachteil, denn mit dieser Methode kann, selbst mit vielen implantierten Multimikroelektroden, immer nur eine Hirnregion auf einmal untersucht werden. Sie ergibt kein Bild der Aktivierungen im gesamten Gehirn, wie beispielsweise fMRT. Für die Untersuchung anderer, möglicherweise korrelierender, Gehirnregionen, muss ein Tier neu präpariert und gegebenenfalls neu trainiert werden. Auf diese Weise schreitet der Erkenntnisgewinn nur langsam voran. (Ruff et al. 2014, 81).

Nachdem wir nun einzelne (Mess-)Techniken, ihre Funktionsweisen sowie Vor- und Nachteile kennengelernt haben, werden wir im nun folgenden Kapitel 5 ihre Anwendung in neuroökonomischen Experimenten betrachten. Dafür werden wir uns auf ein Beispiexperiment mit menschlichen Probanden in einer fMRT-Studie konzentrieren und sehen, wie solche Experimente aufgebaut werden und was es dabei zu beachten gilt. Generell führt das Kapitel 5 nach der getrennten Betrachtung von Verhaltensökonomie und kognitiven Neurowissenschaften in diesem und dem vorherigen Kapitel diese beiden Enden nun zusammen. Es zeigt auf, wie die beiden Wissenschaften zusammenarbeiten und was sie von dieser Zusammenarbeit erwarten können.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Die neuroökonomische Zusammenarbeit

5

In den Kapiteln 3 und 4 haben wir die Verhaltensökonomie und die kognitiven Neurowissenschaften einzeln betrachtet. In diesem Kapitel werden wir nun sehen, wie die beiden Wissenschaften in der Neuroökonomie zusammenarbeiten. Welchen Beitrag bringen die Verhaltensökonominnen ein, welchen die kognitiven Neurowissenschaftler? Wie arbeiten sie zusammen, wie profitieren sie gegenseitig? Diese Fragen werden wir, wieder getrennt nach den Ursprungswissenschaften, in den folgenden beiden Unterkapiteln betrachten. Wir werden sehen, dass die kognitiven Neurowissenschaften aus der ökonomischen Erfahrung mit behavioralen Experimenten lernen können. Gleichzeitig sollten Verhaltensökonominnen Erkenntnisse aus neurowissenschaftlichen Studien nicht als wenig hilfreich verwerfen. Durch den Einsatz moderner Messtechniken können Erkenntnisse entstehen, die in der Ökonomie Forschungsfragen inspirieren und lange Umwege über mehrere behaviorale Experimente verkürzen können.

In Abschnitt 5.1 geht es um die Frage, welche Ideen, Modelle und Arbeitsweisen aus der Verhaltensökonomie in der Neuroökonomie Verwendung finden. Wir werden am Beispiel des *Balloon Analog Risk Task* (BART) sehen, wie psychologische Experimente zur Risikobereitschaft von Individuen in kognitiv neurowissenschaftliche übertragen werden und welche Nachteile das haben kann. Dem gegenüber stellen wir ein Experiment aus der Verhaltensökonomie über faires Verhalten von Individuen, das ein sehr gutes Beispiel dafür ist, wie schwer definierbare Präferenzen mit klaren Benchmarks operationalisiert werden können. Darin liegt der Beitrag der Verhaltensökonomie zur Neuroökonomie begründet.

Abschnitt 5.2 zeigt dann auf, wie die Arbeit der kognitiven Neurowissenschaftler Erkenntnisse für die Neuroökonomie liefert. Hierfür betrachten wir ein beispielhaftes Experiment von Zaki und Mitchell (Zaki et al. 2010) aus der neurowissenschaftlichen Seite der Neuroökonomie, anhand dessen wir nicht nur sehen werden, wie neurowissenschaftliche Experimente mit fMRT-Scannern geplant,

durchgeführt und analysiert werden. Es ist auch ein gutes Beispiel dafür, wie neurowissenschaftliche Erkenntnisse die Verhaltensökonomie bereichern könnte. Zum besseren Verständnis der Arbeit von Zaki und Mitchell und als Grundlage für die Betrachtung von Kritiken in Kapitel 6 werden wir an dieser Stelle auch generell über die Prozesse, Standards und Regularien sprechen, die nötig sind, um fMRT-Experimente erfolgreich durchzuführen.

Das Abschnitt 5.3 fügt die Informationen zusammen. Wie nutzen die Verhaltensökonomie beziehungsweise die kognitiven Neurowissenschaften die Vorteile der jeweils anderen? Wie sieht nun konkret die Zusammenarbeit aus, wie sind sich Neurowissenschaften und Verhaltensökonomie konkret gegenseitig von Nutzen? Die Bilanz kann, wie wir sehen werden, je nach Blickwinkel anders ausfallen.

5.1 Der Beitrag der Verhaltensökonomie zur Neuroökonomie

In Kapitel 3 haben wir uns mit der Entwicklung der Verhaltensökonomie im 20. Jahrhundert beschäftigt. Wir haben gesehen, dass Verhaltensökonomien im Verlauf der vergangenen mehr als fünfzig Jahre Erfahrungen mit Laborexperimenten zu ökonomischen Fragestellungen gesammelt haben. Sie haben begonnen, Experimente in ihren Arbeitsalltag zu integrieren und sie als nützliche Datenquelle zu verwenden. Ihre Fähigkeiten im Umgang mit Experimenten sind dabei stetig gewachsen. Experimentaufbauten, die Komplexität der Aufgabenstellung und die Breite der erforschten Fragen haben sich seit Smiths und Seltens ersten Versuchen ebenso stark verändert wie die Aufbereitung und Analyse der generierten Daten. Betrachtet man Smiths und vor allem Seltens erste Experimente, so fällt deren Komplexität auf. Beispielsweise in dem oben beschriebenen Experiment von Selten sehen sich die Probanden zeitlich ausgedehnten Experimenten gegenüber, für die hohe Konzentration erforderlich ist, denn auch die Aufgabenstellungen sind vielfältig. Nebenbei wird nicht nur ein Protokoll der Entscheidungen und Diskussionen erwartet, sondern auch die eigenständige Skizzierung verschiedener Kosten- und Preis-Absatz-Kurven.

In dem Experiment wird nicht unbedingt eine einzige Hypothese getestet, sondern eher eine spannende Marktsituation erzeugt, in der das Verhalten der Probanden beobachtet wird, um eventuelle Verhaltensmuster zu erkennen, für die dann eine Erklärung gesucht wird. Diese Art des Experimentierens ist vermutlich auch darauf zurückzuführen, dass die experimentelle Ökonomie noch nicht etabliert war und noch wenig Erfahrungen darüber gemacht worden waren, welche

Komplexität an Aufgaben für welchen Zweck angemessen sei, oder grundsätzlich welche Vielzahl an Möglichkeiten sich mit der Durchführung von Experimenten eröffneten. (Abbink et al. 2010, 53).

Heute ist diese Art zu experimentieren etwas aus der Mode gekommen und wurde von Experimenten abgelöst, die darauf abzielen, bestimmte, vorher präzise gestellte Forschungsfragen zu beantworten. Die Experimente werden dann so gestaltet, dass sie genau diese Forschungsfragen adressieren. Dabei ist es eher zweitrangig, ob die Experimentsituation besonders spannend ist, oder genau der Realität entspricht. (Abbink et al. 2010, 54).

Auch die Art der Protokollierung, der Datenaufbereitung, -analyse und -archivierung hat sich durch den Einzug von Computern und entsprechender Software in die Labore der Ökonomen stark verändert. Heute sind die Probanden nicht mehr über Telefone vernetzt, sondern über die vor ihnen stehenden Computer und auch Diagramme zur Datenveranschaulichung werden nicht mehr mit dem Lineal auf Millimeterpapier gezeichnet.

Die Verhaltensökonomien haben Erfahrungen gesammelt und ihre Experimente weiterentwickelt. Unzählige Lehrbücher weisen den Weg zum passenden Experiment für unterschiedliche Zwecke. Die Verhaltensökonomien haben viel Licht in den dunklen Raum der sie umgebenden Möglichkeiten gebracht. Gleichzeitig haben sie die Spieltheorie als hilfreichen Bestandteil der Experimentalanalyse gefestigt.

Die Verhaltensökonomien verfügen heute also über jahrelange Erfahrung im Umgang mit klar definierten und adäquat konstruierten Experimenten und deren Analyse. Dieses Wissen macht sich die Neuroökonomie zunutze. Innerhalb der Neurowissenschaften sind die kognitiven Neurowissenschaften ein noch relativ junges Feld. Ihre experimentellen Anforderungen weichen zum Teil stark von denen anderer neurowissenschaftlicher Felder ab. Die Untersuchung einzelner Neuronen im Bein einer Heuschrecke ist reichlich verschieden von der Untersuchung mehrerer Zellen beziehungsweise des gesamten menschlichen Gehirns. Daher standen die kognitiven Neurowissenschaftler vor ungefähr fünfzehn Jahren vor einem Problem: Sie wollten Experimente durchführen, aber ihnen fehlten Erfahrung und Geräte. Die Ausstattung mit Messinstrumenten ist in den vergangenen Jahren stark gestiegen, die Qualität von Elektroenzephalografen bis fMRT-Scannern hat sich stark verbessert, ebenso ihre Verfügbarkeit. Besonders Elektroenzephalografie erfreut sich aufgrund geringer Kosten und relativ einfacher Anwendung einiger Beliebtheit.

Was den Aufbau, die Durchführung und Analyse von Experimenten angeht, wandten sich die kognitiven Neurowissenschaftler den Psychologen zu. Deren Erfahrung mit Experimenten über menschliches Verhalten war ihnen hilfreich

beim Aufbau eigener Experimente. Zudem besaßen die Psychologen bereits Modelle zur Beschreibung dessen, was sich im Kopf eines Menschen abspielt, wenn er eine Entscheidung trifft und hatten auch passende Experimente und Analysetechniken entwickelt. Diese wurden von den frühen kognitiven Neurowissenschaftlern gern in eigenen Experimenten implementiert. Dieses Übernehmen von Ideen und Vorgehensweisen hatte für die Neurowissenschaftler Vorzüge, brachte aber auch die Schwächen psychologischer Experimentaufbauten mit. Da dies wichtig ist, um den Beitrag der Verhaltensökonomie zur Neuroökonomie zu verstehen, betrachten wir dazu ein Beispiel aus der Psychologie, das einerseits anschaulich aufzeigt, wie psychologische Experimentparadigmen in neurowissenschaftliche fMRT-Experimente umgestaltet werden können. Andererseits zeigt es auch die Nachteile, die auftreten, wenn sich kognitive Neurowissenschaftler ausschließlich auf Konzepte aus der Psychologie verlassen. Unser Beispiel dafür ist der *Balloon Analog Risk Task* (BART), ein Experiment zur Untersuchung des Risikoverhaltens von Individuen. Es wurde von einem Team um den klinischen Psychologen Carl Lejuez (Lejuez et al. 2002) entwickelt und 2002 veröffentlicht als ein Werkzeug zur Untersuchung von Risikobereitschaft (*Risk Taking*). Das Experiment ging der Frage nach, wie sich Individuen bei Entscheidungen unter Risiko verhalten. Im Jahr 2008 nutzten der Neurowissenschaftler Hengyi Rao (Rao et al. 2008) und sein Team den BART als eine der Ersten für ein fMRT-Experiment. Im BART sitzen die Probanden vor einem Bildschirm, der einen Ballon zeigt. Zu Beginn des Experiments ist er nicht aufgeblasen. Die Probanden können nun zwischen zwei Handlungen wählen: Sie drücken den rechten Knopf, der den Ballon ein Stück weit aufpumpt. Dadurch erhalten sie einen Geldbetrag auf einem virtuellen Konto gutgeschrieben. Mit jedem weiteren Aufpumpen erhöht sich der Kontostand, aber auch das Risiko, dass der Ballon platzt. Platzt der Ballon, ist der bei diesem Ballon gewonnene Betrag verloren und wird vom Konto des Probanden wieder abgezogen¹. Nach jedem Mal Aufpumpen gibt es eine kurze Sperre von 1,5–2,5 Sekunden (vom Computer zufällig bestimmt), in der ein kleines rotes Licht am Bildschirmrand leuchtet. Nach dieser Zeit leuchtet das Licht wieder grün und die Probanden können entscheiden, ob sie mit einem Druck des rechten Knopfes den Ballon weiter aufpumpen oder mit einem Druck auf den linken Knopf die Runde beenden möchten. In diesem Fall erscheint der

¹ Die Beschreibung gilt für den BART im fMRT-Experiment von Rao et al. Sie ist an diesem Punkt leicht abgewandelt von der Originalversion von Lejuez et al., wo es ein temporäres sowie ein permanentes Konto gibt. Rao et al. erhöhen den Strafcharakter durch einen Abzug des bereits erhaltenen Betrags vom Konto, während bei Lejuez et al. das permanente Konto sicher ist und unangetastet bleibt. Die Unterschiede sind in unserem Kontext nicht relevant, können jedoch bei (Rao et al. 2008) sowie bei (Lejuez et al. 2002) nachgelesen werden.

Schriftzug „*You Win!*“ auf dem Bildschirm und der bisher gesammelte Geldbetrag bleibt dem Probanden erhalten. Mit jedem weiteren Aufpumpen des Ballons erhalten die Probanden mehr Geld. Das Risiko, dass der Ballon beim nächsten Druck auf den rechten Knopf platzt, erhöht sich dabei auch stetig. Dieser Anstieg des Risikos war den Probanden bekannt, nicht aber, welche Wahrscheinlichkeit zu platzen bei welchem Schritt genau vorlag. An einem vom Computer bestimmten Punkt, der bei jedem Ballon verschieden ist, platzt der Ballon. In diesem Fall erscheint der Text „*You Lose!*“ und der mit diesem Ballon gewonnene Geldbetrag wird vom Konto des Probanden abgezogen. Jeder Ballon kann maximal zwölf Mal aufgepumpt werden, dann ist der maximale Geldbetrag erreicht – sofern der Ballon nicht vorher zerplatzt ist. Die Versuchszeit wurde von Rao et al. (Rao et al. 2008, 905) auf acht Minuten festgelegt. Die Anzahl der Ballons, die die Probanden in diesem Zeitraum zu bearbeiten hatten, war dagegen nicht begrenzt, sondern ermaß sich an der Geschwindigkeit, mit der die Probanden ihre Entscheidungen trafen. Wer schneller den rechten oder linken Knopf drückte, konnte mehr Ballons in den acht Versuchsminuten unterbringen. Im fMRT-Scanner liegend konnten die Probanden über einen Spiegel das Geschehen auf dem Bildschirm betrachten, der hinter der magnetischen Röhre angebracht war. Für ihre Entscheidungen bekamen die Probanden eine Box mit zwei Knöpfen darauf in den Händen, die sie mit dem linken beziehungsweise dem rechten Daumen betätigen konnten.

Um vergleichen zu können, ließen Rao et al. ihre Probanden nach dem ersten, beschriebenen Durchgang einen zweiten, leicht modifizierten Durchgang von ebenfalls acht Minuten durchlaufen, um ein passives Eingehen von Risiken ohne eigene aktive Entscheidung zu simulieren. In diesem zweiten Versuchsdurchlauf waren keine Wahlmöglichkeiten für die Probanden gegeben, sie sollten die Luftballons einfach immer weiter aufpumpen. Der Computer entschied in diesem Durchlauf darüber, ob der Ballon platzte und der Gewinn verloren war, oder ob er seine endgültige Größe erreichen und der Proband seinen Gewinn erhalten würde. Die Probanden hatten hier also keinerlei Wahlmöglichkeiten, sie konnten nicht aussteigen und den bisherigen Gewinn behalten.

Ein kurzes Wort zu den Ergebnissen: Rao et al. (Rao et al. 2008, 907) beobachteten, dass das Verhalten der Probanden in beiden Versuchsdurchläufen, dem aktiven und dem passiven, ähnlich war. Im aktiven Durchlauf bearbeiteten sie 20.1 Ballons, im passiven 22.1, wobei sie jeden Ballon 7,6 beziehungsweise 7,5 Mal aufpumpten. Es fand sich keine signifikante Verbindung zwischen der Art des Versuchsdurchlaufs (aktiv mit Wahl oder passiv ohne Wahl) und Ergebniswert. Es fanden sich keine signifikanten Unterschiede in der Reaktionszeit zwischen den beiden Versuchsdurchläufen. Die Ergebnisse über die Gehirnaktivitäten aus dem fMRT-Scanner waren etwas aufschlussreicher für die Forscher,

da sie die von ihnen im Vorfeld vermuteten Gehirnregionen aktiviert sahen, auch wenn die Unterschiede erst unter Nutzung eines geringeren Schwellenwertes für signifikante Neuronenaktivität ersichtlich wurden als zuvor angedacht.

Der *Balloon Analog Risk Task* ist ein interessantes Experiment, das dazu konstruiert wurde, die Risikobereitschaft von Menschen zu untersuchen. Allerdings birgt das Experiment auch Nachteile. Vor allem stellt sich bei der Analyse der Daten die Frage, wie sie zu interpretieren sind. Welche Werte kennzeichnen beispielsweise einen besonders risikofreudigen Probanden oder einen besonders risikoaversen? Wie oft Aufpumpen pro Ballon spricht für eine risiko-freudige Testperson, wie oft für eine risiko-averse? Wann haben die Probanden *richtig* gehandelt und welche Werte sind abnorm? Der *Balloon Task* scheint hier etwas willkürlich zu sein, was für psychologische Modelle nicht ungewöhnlich sein muss, da sie nicht selten auf Konstruktbasis angelegt sind. „Konstrukte können nicht beobachtet, sondern müssen aus Beobachtung erschlossen werden.“, wie Meindl (Meindl 2011, 5) auf den Punkt bringt. Das ist ein systematischer Nachteil, der nicht so einfach aus der Welt zu schaffen ist und sich auch auf die Auswertung kognitiver neurowissenschaftlicher Experimente auswirkt. Betrachten wir im Gegensatz dazu nun ein Beispiel, wie Verhaltensökonomien schwer zu definierende Präferenzen von Individuen operationalisieren.

Im Vergleich mit den gerade betrachteten psychologischen Modellen sehen Ökonomen ihre Modelle im Vorteil: Mit ihren Konzepten von Nutzen und *revealed Preferences* haben sie ein Modell an der Hand, das grundsätzlich auf jede Situation anwendbar ist. Im Gegensatz zu vielen psychologischen Experimenten bieten ökonomische Experimente klar strukturierte Situationen mit gewissen Wahrscheinlichkeiten. Es ist klar, was bekannt ist und was nicht, was *richtig* ist und was nicht. Es ist über den Versuchsaufbau bereits klar definiert, welches Verhalten welche Schlussfolgerungen zulässt. Die klare Interpretierbarkeit zeigt sich beispielsweise anschaulich an Experimenten, die Fairness (soziale Präferenzen) untersuchen, wie zum Beispiel *Gift-Exchange-Experimente*. Das *Gift-Exchange-Spiel* ist eine Form von sequenziellem Gefangenendilemma und zeichnet sich dadurch aus, dass die Probanden die Möglichkeit haben, Geld auszugeben, um einer anderen Person zu helfen, in der Hoffnung, dass diese Person ebenfalls Kosten auf sich nehmen wird, um die Hilfe zurückzugeben. Hier geht es also um Kooperation und welche Kosten die Probanden dafür bereit sind, auf sich zu nehmen. Eine Schwierigkeit bei einem solchen Experiment ist die Beurteilung der Nettigkeit oder Fairness der Probanden. Fairness ist ein schwer zu definierender Begriff. Wann verhält sich eine Person fair, wann unfair? Es scheint schwierig, einen Begriff wie diesen für ein Experimentparadigma zu operationalisieren.

Hier zeigen verhaltensökonomische Modelle ihre Stärke, und deshalb ist das *Gift-Exchange*-Spiel auch so ein gutes Beispiel: Anders als bei der Bewertung der Risikoneigung im Ballonexperiment, ist im *Gift-Exchange*-Spiel klar, dass diejenigen Probanden Fairness in ihr Handeln einbeziehen, die Kosten für sich selbst aufnehmen, die nicht unbedingt notwendig sind für die ökonomische Transaktion im Spiel. Auch über den Grad der Berücksichtigung von Fairness-Überlegungen lässt sich etwas aussagen. Betrachten wir zur besseren Veranschaulichung das erste veröffentlichte *Gift-Exchange*-Spiel, das von Ernst Fehr und seinem Team stammt (Fehr et al. 1993):

Die Experimentatoren wollten Verhalten auf dem Arbeitsmarkt untersuchen, hatten jedoch Bedenken, dass die sozialen Präferenzen der Probanden eine größere Rolle spielen könnten, wenn sie über dieses Ziel der Studie informiert würden, da das Experimentatorenteam soziale Interaktion zwischen Unternehmer und Mitarbeitern als intensiver einschätzten als die zwischen Käufern und Verkäufern auf einem Gütermarkt. Sie maskierten also für die Probanden die Instruktionen als Instruktionen für einen Markt, auf dem ein bestimmtes Konsumgut gehandelt wird. Die Probanden wurden per Los aufgeteilt in Käufer (Arbeitgeber) und Verkäufer (Arbeitnehmer) dieses Guts.² Dabei wurden mehr Arbeitnehmer als Arbeitgeber ausgelost, um eine Konkurrenzsituation unter den Arbeitern zu schaffen. Das Experiment bestand aus zwei Phasen. In Phase 1, die auf 3 Minuten begrenzt war, machten die Arbeitgeber Lohnangebote an die Arbeitnehmer. Diese konnten die Angebote annehmen. Wurde ein Angebot akzeptiert, kam sogleich ein Vertragsverhältnis zwischen Arbeitgeber und Arbeiter zustande, das beide vom Markt nahm und für sie Phase 2 des Experimentes eröffnete. In dieser Phase legten nun die Arbeiter den Grad an Anstrengung oder Mühe fest, mit dem sie ihre Arbeit ausführen würden. Diese Antwort wurde nur dem jeweiligen Arbeitgeber offenbart, die übrigen Versuchsteilnehmer erfuhren diese Antworten nicht. Um die gegenseitige Anonymität während des Versuchs zu sichern, saßen die Arbeitgeber und Arbeitnehmer getrennt voneinander in unterschiedlichen Räumen. Die Kommunikation zwischen den beiden Gruppen verlief über zwei Experimentatoren, die je in einem Raum saßen und über Telefon verbunden waren. Sie gaben ausgesprochene Angebote aus dem Arbeitgeberraum an den Arbeitnehmerraum weiter, wo sie öffentlich ausgerufen wurden. Akzeptierte Angebote wurden ebenso zurückübermittelt.

² Wie auch Fehr et al. verwende ich hier das Arbeitsmarktvokabular. Die Probanden erhielten Instruktionen, auf denen nur Begriffe eines Gütermarktes verwendet wurden, die zu den Arbeitsmarkt Begriffen analog verwendbar sind. (Fehr et al. 1993, 439).

Ein Ziel des Experiments war es, den Einfluss von Fairness auf die Marktpreise zu untersuchen. Um die Versuchsdaten später dahingehend interpretieren zu können, muss klar sein, was Fairness bedeutet, wann ein Proband fair gehandelt hat und wann nicht. Fehr und sein Team bauten ihren Versuch daher so auf, dass Fairnessparameter eindeutig zu erkennen sind (Fehr et al. 1993, 441–442):

„Although it is difficult to give an exact definition of fairness, the term necessarily involves some comparison between the gains of transactors. Therefore, if an agent is motivated by fairness considerations, his actions are based not only on his own gains, but also on the gains of other parties. Our experiment aimed at detecting whether agents exhibit intrinsic fairness preferences, and whether they base their actions on the expectation that others act fairly.“

Fairness bedeutet also, dass die Arbeitgeber und Arbeitnehmer im Experiment nicht nur ihren eigenen, sondern auch den Gewinn ihres Gegenübers im Blick haben. Dazu muss gewährleistet sein, dass sich die Probanden im Klaren darüber sind, wann sie mehr Geld ausgeben, als nötig ist, um die Transaktion durchführen zu können. Daher erhielten sie Informationen über ihre Kostenfunktionen. Alle Versuchsteilnehmer, egal, welcher Gruppe sie angehörten, wussten, dass es mehr Arbeiter als Arbeitgeber gab und daher ein Konkurrenzdruck unter den Arbeitern vorhanden war (wer keinen Vertrag in Phase 1 abschließen konnte, erhielt keinen Geldbetrag für diese Periode). Die Arbeitgeber waren sich also im Klaren darüber, dass die Arbeiter im Grunde jedes Angebot oberhalb des Gleichgewichtslohnes akzeptieren würden (Fehr et al. 1993, 450). Gleichzeitig waren sich die Arbeitnehmer im Klaren darüber, dass die Arbeitgeber per Experimentdesign keine Verluste erwirtschaften konnten (siehe weiter unten Erläuterungen zu (2)). Desweiteren waren allen Teilnehmern die Auszahlungsfunktionen der beiden Versuchsgruppen bekannt (Fehr et al. 1993, 441):

Die Auszahlung der Arbeiter war gegeben durch

$$(1) u_j = p_j - c - m(e_j).$$

c steht für die Kosten für die Leistung von 1 Einheit Arbeit durch Arbeiter j und war in den Versuchen für alle Arbeiter auf 26 festgelegt. p_j gibt den Lohn, den der Arbeiter vom Arbeitgeber erhält und e_j die angewandte Mühe des Arbeiters. $m(e_j)$ übersetzt diesen Wert in monetäre Kosten, die dem Arbeiter durch das Aufwenden von Mühe entstehen. Diese Kosten steigen stetig und sind auf einer Tabelle festgehalten, die den Versuchsteilnehmern ausgeteilt wurde. Zusätzlich gilt: $m(e_{min}) = 0$. Die Auszahlung eines Arbeiters hängt also nicht nur von der Höhe des Lohnes ab, den er in Phase 1 akzeptiert hat, sondern auch von der Höhe der Mühe, mit der er seine Arbeit verrichtet. Grundsätzlich muss er nicht

mehr als e_{min} für seine Arbeit aufwenden, was seine Auszahlung am größten halten würde. Allerdings hat der Grad seiner Mühe auch Auswirkungen auf die Auszahlung des Arbeitgebers i :

$$(2) \Pi_i = (v - p_i)e_i.$$

Das Einkommen des Arbeitgebers wird durch ve_i bestimmt. Eine Einheit an Mühe, die der Arbeiter aufwendet, produziert eine Anzahl von v Einheiten eines Produkts, das für einen Preis von 1 verkauft wird. Die Annahme, dass die Lohnkosten p_i ebenfalls von der Mühe des Arbeiters abhängen, soll verhindern, dass der Arbeitgeber Verluste erwirtschaften kann. Die Anzahl produzierbarer Einheiten wurde auf $v = 126$ festgelegt.

Die Arbeitgeber haben durch den unter den Arbeitnehmern herrschenden Konkurrenzdruck keine Notwendigkeit, mehr als den Gleichgewichtslohn zu zahlen, denn die Arbeiter sind an einem Vertragsabschluss sehr interessiert und müssen auf eingehende Angebote blitzschnell reagieren, da sie ansonsten leer ausgehen. Sogar Angebote unterhalb des durchschnittlichen Niveaus wurden durchweg akzeptiert.³ Gleichzeitig haben die Arbeiter, die ihren Lohn bereits in Phase 1 erhalten haben, keine Notwendigkeit, in Phase 2 eine höhere Mühe als e_{min} einzusetzen. Diese Verhaltensweisen würden einem rationalen Vorgehen entsprechen, vor allem da die Arbeitgeber und Arbeitnehmer gegenseitig anonym bleiben. Die Arbeiter müssen also nicht befürchten, aufgrund minimaler Leistung in dieser Runde, in späteren Runden niedrigere Angebote zu erhalten. Die Arbeitgeber können aber mehr Lohn anbieten, um die Motivation der Arbeitnehmer zu erhöhen. Diese *revanchieren* sich dann mit mehr Mühe, die auch die Auszahlung des Arbeitgebers erhöht. Die Arbeitgeber können also fair handeln (mehr Lohn bezahlen als der Gleichgewichtslohn) in der Hoffnung, dass auch der Geschäftspartner ihnen gegenüber fair handelt (mehr Mühe aufwenden als e_{min}). Im Experiment von Fehr und seinem Team handeln die Teilnehmer auch beinahe ausnahmslos so. Die Arbeitnehmer wählten alle möglichen Level von Mühe, von 0 bis 1, wobei der Minimalwert in 44 von 276 Fällen gewählt wurde, der Maximalwert in 6 Fällen. Damit war der durchschnittlich gewählte Wert von 0,4 viermal so hoch wie von der Standardtheorie angenommen worden wäre. Fehr und seine Kollegen beobachteten, dass sowohl Durchschnitt als auch Median der von den Arbeitnehmern gewählten Mühe mit dem Gehalt anstieg. Die Arbeitgeber zahlten im Durchschnitt mehr als den Gleichgewichtslohn, es wurde auch keine

³ Fehr et al. berichten (Fehr et al. 1993, 451), dass die Angebote innerhalb von Sekundenbruchteilen akzeptiert wurden. In nach dem Experiment ausgefüllten Fragebögen antworteten arbeitslos gebliebene Arbeitnehmer, dass sie deshalb kein Angebot angenommen hatten, weil andere schneller gewesen waren, und nicht etwa, weil sie nicht gewollt hätten.

Konvergenz des Durchschnittslohns zum von der Standardtheorie vorhergesagten Gleichgewichtslohn beobachtet.

Wichtig ist bei der Analyse aber vor allem, dass es in diesem Experiment ein klares Kriterium dafür gibt, was Nettigkeit oder Fairness bedeutet und wann ein Versuchsteilnehmer fair gehandelt hat. Bietet ein Arbeitgeber mehr Lohn als der Gleichgewichtslohn, beziehungsweise bringt ein Arbeitnehmer mehr Mühe auf als e_{min} , dann war das faire Handeln, ein klares Ja-/Nein-Kriterium. Zudem lässt sich an der genauen Höhe des Lohnangebots, beziehungsweise der eingesetzten Mühe, ein gradueller Unterschied in der Fairness ausmachen.

Ein solcher Aufbau findet sich nicht nur im *Gift-Exchange-Experiment*, sondern ist charakteristisch für die meisten ökonomischen Experimente. Dadurch bringt ein ökonomisches Experiment leichter klare Vergleichskriterien hervor als ein Experiment wie der *Balloon Analogue Risk Task*, was nur von Vorteil sein kann, wenn man das Verhalten verschiedener Menschen in verschiedenen Situationen untersucht. Zudem sind die ökonomischen Strategien, einen Versuch aufzubauen, auf sehr viele Situationen anwendbar.

Natürlich können kognitive Neurowissenschaftler ihre Forschungsarbeit ohne Anleihen aus dem ökonomischen Labor durchführen. Natürlich können sie Modelle selbst erstellen oder sie in anderen Wissenschaften suchen. Allerdings verzichten sie in diesem Fall auf die Vorteile, die ökonomische Modelle mit sich bringen würden. Genau das ist der Beitrag, den Ökonomen zur Neuroökonomie beitragen. Sie bringen die inzwischen Jahrzehnte lange Erfahrung ihres Fachgebietes im Bereich der Entscheidungstheorie mit, ebenso wie ihre Erfahrung, daraus aussagekräftige Experimente aufzubauen und auszuwerten.

Im folgenden Abschnitt 5.2 werden wir sehen, wie die Anwendung verhaltensökonomischer Modelle im neurowissenschaftlichen Experiment aussehen kann und welche Möglichkeiten sich dadurch eröffnen.

5.2 Der Beitrag der kognitiven Neurowissenschaften zur Neuroökonomie

Der Beitrag der kognitiven Neurowissenschaften zur Neuroökonomie besteht zu einem großen Teil aus der empirischen Generierung von Daten. In unzähligen verschiedenen Studien untersuchen kognitive Neurowissenschaftler unterschiedliche Fragestellungen. Die dabei am häufigsten eingesetzten empirischen Methoden haben wir bereits in Abschnitt 4.3 kennengelernt. In den Studien geht es darum, Aktivitätsmuster einzelner Neuronen, einzelner Hirnregionen oder des gesamten Gehirns zu erfassen und mit einem bestimmten Verhalten in Verbindung

zu bringen. Viele gesammelte Daten sollen dann Rückschlüsse darauf zulassen, wie das Gehirn in welcher Situation arbeitet. Klassische Experimente aus den kognitiven Neurowissenschaften sehen beispielsweise Erinnerungsaufgaben für die Probanden vor, während derer ihre Gehirnaktivitäten mit unterschiedlichen Hilfsmitteln erfasst und aufgezeichnet werden. Aus vielen solcher Daten kann dann ein Modell der Arbeit des Gedächtnisses entstehen. In der neuroökonomischen Forschung läuft der Erkenntnisgewinn ähnlich, hier findet aber der in Abschnitt 5.1 besprochene Beitrag der Verhaltensökonomie Anwendung: Der Aufbau der Studien und der grundsätzliche Ablauf der Experimente ist gleich oder sehr ähnlich wie bei klassisch kognitiv neurowissenschaftlichen Studien, es werden aber verhaltensökonomische Fragestellungen und Methoden zur Analyse der verhaltensbezogenen Daten zugrunde gelegt. Beispielsweise spielen Versuchspersonen in einer bildgebenden Röhre oder unter einer EEG-Kappe ökonomische Spiele. Das kann in Gruppen geschehen oder einzeln, je nach Versuchsaufbau. Allerdings ist es keine triviale Aufgabe, ein ökonomisches Spiel wie das Vertrauensspiel (*Trust Game*) in ein im Scanner spielbares und am Ende analysierbares Verfahren umzuwandeln. Die Vorbereitungen für Studien können Monate dauern.

Um den Beitrag der kognitiven Neurowissenschaften zur Neuroökonomie anschaulich zu machen und besser zu verstehen, werden wir in diesem Kapitel den grundlegenden Ablauf neuroökonomischer Studien generell betrachten. Wie wir sehen werden, gibt es bei der Planung, Durchführung und Analyse neurowissenschaftlicher Studien zahlreiche Einflussgrößen zu beachten. Da sich Studien und Abläufe je nach Zielsetzung und Hilfsmittel unterscheiden können, an dieser Stelle aber leider nicht der Platz für eine vollständige Darstellung aller in Frage kommenden empirischen Methoden ist, nehmen wir Studien mit funktionseller Magnetresonanztomographie (fMRT) als Beispiel. Mit der Entwicklung der bildgebenden Verfahren Ende der 90er Jahre eröffnete sich den Neurowissenschaftlern eine ganze Fülle neuer Möglichkeiten zur Erforschung menschlicher Gehirnaktivitäten. fMRTs sind seither zu einem beliebten Instrument für neurowissenschaftliche und neuroökonomische Untersuchungen geworden, weshalb wir mit der Betrachtung von fMRT-Studien einen guten Einblick in die tägliche Arbeit der kognitiven Neurowissenschaftler erhalten. In diesem Kapitel werden wir anhand eines Beispielerperiments von Zaki und Mitchell (Zaki et al. 2010) sehen, wie neuroökonomische Studien mit bildgebenden Verfahren aussehen können. Wir werden die verschiedenen Stufen von Studien betrachten, von der Planung über die Durchführung bis zur Analyse der gesammelten Daten. Um die Aktionen Zakis und Mitchells in unserem Beispielerperiment besser einordnen zu können, werden wir uns nicht nur auf ihre Entscheidungen in dieser konkreten Studie konzentrieren, sondern auch generelle Ideen, Regelungen und

Empfehlungen aus den Neurowissenschaften hinzuziehen. Das wird dabei helfen, zu verstehen, wie neurowissenschaftliche Studien aufgebaut sind und mit welchen Schwierigkeiten Neurowissenschaftler bei ihrer Arbeit umgehen müssen, was wiederum dabei hilft, den Beitrag der kognitiven Neurowissenschaften zur Neuroökonomie zu verstehen. Zudem bereitet dieses Wissen den Grund für die Betrachtung einiger Kritiken in Kapitel 6.

Wie auch bei ökonomischen oder anderen Studien müssen bei der Planung von fMRT-Studien einige Dinge bedacht werden, bevor das tatsächliche Experiment durchgeführt werden kann. Allerdings ist die Nutzungszeit für fMRT-Geräte relativ teuer, sodass nicht nur ganze Studien sondern auch einzelne Scanner-Durchgänge nicht ohne Weiteres wiederholt werden können. Gerade wegen der hohen Kosten ist nicht jeder neurowissenschaftliche Fachbereich mit einem fMRT-Gerät ausgestattet. Wie bereits in Abschnitt 4.3 erwähnt, nutzen Neurowissenschaftler daher oft Geräte einer nahegelegenen Klinik oder einer anderen neurowissenschaftlichen Forschungseinrichtung, die über einen Scanner verfügt. In beiden Fällen müssen Scannerstunden gemietet werden. Vor allem Krankenhäuser benötigen ihre Geräte tagsüber meist selbst, sodass Experimente nicht selten nachts durchgeführt werden. Eine Gerätestunde kann dabei zwischen \$100 und \$1000 kosten, das richtet sich nach der Höhe der Operationskosten, die die Klinik oder die Forschungseinrichtung für ihren Scanner ansetzen (Carter et al. 2010, 25). Studien sollten also schon allein aus finanzieller Sicht sorgfältig geplant und durchgeführt werden.

Gewöhnlich beginnt die Planung mit der Formulierung einer geeigneten Forschungshypothese. Auch in den kognitiven Neurowissenschaften wird der Effekt einer unabhängigen Variable auf eine abhängige Variable untersucht. Bei fMRT-Experimenten kann die unabhängige Variable ein Stimulus sein, eine Aufgabe oder auch ein Unterschied zwischen den verschiedenen Probanden (beispielsweise Alter oder Krankheitsbild). Die abhängige Variable ist bei fMRT-Experimenten die Intensität des BOLD-Signals für einen bestimmten Teil des Gehirns. Ziel eines fMRT-Experiments ist es also, eine unabhängige Variable auf Signalintensität zu testen. Carter et al. (Carter et al. 2010, 26) geben Beispiele für typische Schemata von Hypothesenformulierungen: „For a given stimulus or task, activity in brain region X precedes activity in brain region Y.“, oder „Activity in brain region X is higher under condition Y than condition Z.“ Die Autoren geben zu bedenken, dass sich in diesen Beispielhypothesen die Experimentatoren bereits auf bestimmte Hirnregionen konzentrieren. Das setzt natürlich voraus, dass eine gewisse Vermutung darüber besteht, welche Hirnregionen für diesen bestimmten Aufbau relevant sind. Bildgebende Verfahren können auch dazu genutzt werden,

das gesamte Gehirn zu analysieren, um so erst relevante Hirnregionen zu identifizieren. Allerdings produzieren fMRT-Studien relativ große Datenmengen, sodass eine Analyse des gesamten Gehirns langwierig sein kann. Daher ist es von Vorteil, wenn die Experimentatoren von vornherein eine Idee haben, welche Hirnregionen sich zu analysieren lohnen.

Ist die Forschungshypothese formuliert, gilt es, ein geeignetes experimentelles Paradigma (*Task Paradigm*) zu finden. Mit experimentellen Paradigmen bezeichnen kognitive Neurowissenschaftler die Strategien, nach denen den Probanden im Experiment die Stimuli präsentiert werden (Carter et al. 2010, 26). Ein gutes experimentelles Paradigma zu finden, ist nicht leicht. Dazu muss das zu untersuchende (motorische, sensorische, kognitive etc.) Phänomen in Versuchsbedingungen übersetzt werden. Gute Paradigmen werden daher zum Teil jahrelang von anderen Wissenschaftlern aufgegriffen und mit Variationen immer wieder verwendet. Neurowissenschaftler greifen deshalb oft auf Paradigmen aus vielen Jahren kognitiv-psychologischer Forschung zurück. Allerdings ist es auch hier nicht einfach, sie an die Erfordernisse von fMRT-Studien anzupassen. In der neuroökonomischen Forschung werden gewöhnlich Paradigmen aus der Verhaltensökonomie eingesetzt (siehe Abschnitt 5.1). (Sommer 2010, 243–245). Auch sie werden gemeinsam mit Verhaltensökonomern im Experimentteam in einen praktikablen Versuchsablauf übersetzt. Dabei muss eine ganze Reihe möglicher Fallstricke bedacht werden: Der Ablauf der Experimentaufgabe sollte für die Probanden leicht verständlich sein, sodass sie im Scanner immer wissen, wie ihre Aufgabe aussieht und was sie zu tun haben. Zudem werden Versuchspersonen im Laufe der Zeit unkonzentriert, sie werden müde, und ihre Leistungsfähigkeit lässt nach. Auch Kopfbewegungen sind ein häufiges Problem. Daher sollten Experimente nicht zu lange dauern und nötigenfalls in mehrere Sessions aufgeteilt werden, zwischen denen eine Pause eingelegt werden kann (die Versuchspersonen können sich die Beine vertreten oder zur Toilette gehen). Daneben kann auch ein interessanter Experimentaufbau mit kurzen Interstimulus-Intervallen (ISI, Leerlaufzeiten, in denen kein Stimulus präsentiert und keine Aktion der Versuchspersonen erwartet wird) helfen, die Probanden konzentriert zu halten.

Auch das sogenannte Rauschen, durch das Messungen fehlerhaft werden können, ist ein Problem, für das die Experimentatoren eine Lösung finden müssen: Natürlich sollte ein gutes Experimentdesign eine Analyse zulassen, welche Hirnaktivität zu welchen kognitiven Funktionen gehört. Das Problem ist dabei nicht nur, dass ein Experiment mehrere Funktionen im Gehirn ablaufen lässt, sondern auch, dass immer ein Rauschen auftritt, das die Signale überlagert. Das Rauschen stammt von denjenigen Hirnaktivitäten, die nicht durch die Experimentaufgabe ausgelöst werden, sondern durch andere psychische und physikalische

Vorgänge, die nicht abgeschaltet werden können sowie durch Rauschen im Messprozess selbst. Die durch den Trial verursachten Signale müssen von den durch Rauschen verursachten unterschieden werden können. Das geschieht beispielsweise dadurch, dass das Experiment aus mehreren Sessions des gleichen Ablaufs besteht, sodass möglichst viele Trials BOLD-Signale liefern, deren Intensität dann gemittelt werden kann. Je mehr Daten zur Verfügung stehen, desto besser lassen sich die Schwankungen herausmitteln. (Sommer 2010 247–248).

Wurde über den Umgang mit dem Rauschen entschieden, gilt es, einen Mechanismus zu finden, um die verschiedenen Aktivierungszustände den im Experiment hervorgerufenen Funktionen zuordnen zu können. Dazu gibt es mehrere Stellschrauben bei der Erstellung der Abläufe. Die erste ist die Wahl des passenden Designs. Ganz allgemein werden faktorielle und parametrische Designs unterschieden (Sommer 2010, 250–252). Mit ihnen wird festgelegt, welche experimentellen Bedingungen miteinander verglichen werden sollen. Bei faktoriellen Designs werden die Hirnfunktionen, die man beim Probanden induzieren will, in mehrere Faktoren aufgeteilt. Diese sollen dann in verschiedenen Bedingungen des Experiments auftreten, sodass die Differenzen der Bedingungen Rückschlüsse auf die Aktivierungen durch die einzelnen Faktoren und ihre Interaktionen zulassen. Um das besser verstehen zu können, betrachten wir zur Veranschaulichung kurz ein Beispiel von Sommer, ein Experiment von Karl Friston, bei dem es um phonologischen Abruf geht. Der phonologische Abruf bezeichnet „das Finden des Namens für das bereits erkannte Objekt“ (Sommer 2010, 249). Um diesen zu untersuchen, könnte man nun eine Versuchsperson in einen fMRT-Scanner schieben, ihr ein Bild eines, sagen wir, Stuhls vorsetzen (dieses Objekt erkennt sie ganz bestimmt) und ihre Hirnaktivitäten scannen, und die Publikation ist so gut wie fertig. Allerdings geschehen bis zum phonologischen Abruf noch einige andere Prozesse im Gehirn. Der visuelle Stimulus muss aufgenommen und analysiert und das abgebildete Objekt erkannt werden. Zudem können die Experimentatoren einen tatsächlich stattgefundenen phonologischen Abruf nur dadurch überprüfen, dass die Versuchsperson das Objekt laut bezeichnet, sprachlichen Output gibt es also auch. In den Daten im Computer finden sich also nicht nur die Aktivierungen, die direkt mit dem phonologischen Abruf zusammenhängen, sondern auch die Aktivierungen darum herum. Um das Aktivierungsmuster des phonologischen Abrufs allein herauszurechnen, wird der Versuchsablauf in verschiedene Faktoren unterteilt, die dann in unterschiedlichen Konstellationen auf verschiedene Bedingungen verteilt werden. So wird in Bedingung 1 zu Beginn des Experiments der Versuchsperson ein Bild gezeigt, auf dem noch kein Objekt zu sehen ist, sondern eine Farbe. Wenn die Person das Bild sieht, soll sie „Ja“ sagen. Damit findet sich in Bedingung 1 der Faktor *visuelle*

Analyse und der Faktor *sprachlicher Output*. In Bedingung 2 erscheint auf dem farbigen Bild ein Objekt, in unserem Fall ein Stuhl. Dazu soll die Versuchsperson wieder „Ja“ sagen. In dieser Bedingung kommt also zu den beiden ersten Faktoren (die liegen ja noch immer an) der Faktor *Objekterkennen* hinzu. In Bedingung 3 soll die Versuchsperson nun das Objekt benennen. Hier kommt zu den drei bisherigen Faktoren noch der Faktor *phonologischer Abruf* hinzu. Diesen findet man bisher nur in dieser einen Bedingung, was die Datenlage nicht eindeutig macht, also kommt noch eine vierte Bedingung ins Spiel. In dieser verschwindet das Objekt wieder und nur das Bild mit der Farbe vom Anfang des Experiments ist sichtbar. Nun soll die Versuchsperson die Farbe benennen. So kommen hier die Faktoren *visuelle Analyse*, *phonologischer Abruf* (arbeitet auch für Farben) und *sprachlicher Output* zusammen. In dieser Konstellation findet sich jeder Faktor in mindestens zwei Bedingungen in je anderen Konstellationen. So können die Bedingungen, beziehungsweise die für sie aufgezeichneten Hirnaktivitäten, voneinander abgezogen werden. Auf diese Weise erhält man die Aktivierungsmuster für einzelne Faktoren oder deckt Interaktionen zwischen ihnen auf.

Eine zweite Möglichkeit, den Ablauf eines Experiments zu strukturieren, sind parametrische Designs. Hier wird zwischen den verschiedenen Bedingungen nicht die Aufgabe verändert wie beim Experiment vom phonologischen Abruf, sodass Interaktionseffekte gar nicht erst auftreten können. Stattdessen wird zwischen den Bedingungen die Menge der Prozessierungen variiert. Das heißt, die Bearbeitung der Versuchspersonen erfolgt in unterschiedlichen Abstufungen. Zum Beispiel kann sich bei einer Aufgabe der Schwierigkeitsgrad ändern oder die Stärke der Schmerzen, die den Versuchspersonen zugefügt werden. Dieses Design setzt natürlich voraus, dass die Prozessierungen auch abstufbar sind, der phonologische Abruf wäre es beispielsweise nicht, dafür würde sich ein parametrisches Design schlecht eignen.

Eine dritte Möglichkeit, die nur kurz Erwähnung finden soll, ist die in früheren Zeiten oft verwendete Subtraktionsmethode. Sie wird heute kaum noch eingesetzt und ist ein Beispiel dafür, wie ein schlechtes Experimentdesign das Herausfiltern der aufgabeninduzierten Aktivierungen verhindern kann. Die Idee der Subtraktionsmethode stammt aus dem Jahr 1868. Sie sollte dabei helfen, die Länge eines Verarbeitungsprozesses zu messen. Dabei bearbeitet die Versuchsperson zwei Aufgaben, die sich in nur einem, genau definierten, Verarbeitungsprozess unterscheiden. Die Reaktionszeit beider Aufgaben wird gemessen, wobei die Zunahme an Zeit für die komplexere Aufgabe als die Zeit interpretiert wird, die es braucht, um den zusätzlichen Verarbeitungsprozess zu absolvieren. Diese Überlegung setzt allerdings eine serielle Verarbeitung und eine lineare Zunahme an Aktivitäten im Gehirn voraus und vernachlässigt dabei Interaktionen und Abhängigkeiten bei

der Verarbeitung von Prozessen. (Sommer 2010, 232). Von solchen muss aber vor allem bei komplexeren kognitiven Prozessen ausgegangen werden, sodass die Subtraktionsmethode viel zu simple Annahmen macht, um sie in modernen fMRT-Studien einzusetzen.

Die zweite Stellschraube für die Erstellung eines geeigneten Experimentablaufs in einem fMRT ist die zeitliche Anordnung des Experiments. In welcher Reihenfolge und in welchem zeitlichen Abstand sollen den Versuchspersonen die Stimuli präsentiert werden? Bei dieser Entscheidung spielt der BOLD-Effekt wieder eine Rolle, denn die hämodynamische Antwort, also der Anstieg des Blutflusses zu den aktivierten Nervenzellen, dauert von Beginn bis zum vollständigen Abklingen circa 15 bis 20 Sekunden. Werden die Stimuli in einem kürzeren Abstand präsentiert, addieren sich die hämodynamischen Antworten der verschiedenen Stimuli. Bis zu einem Abstand von 2 Sekunden addieren sie sich linear, also berechenbar, darunter jedoch entstehen komplexe Interaktionen, die die Daten kaum analysierbar machen. Interstimulus-Intervalle von weniger als 2 Sekunden sollten also vermieden werden. Je nach Experimentdesign ist eine lineare Addition der Aktivierungen sogar gewünscht. In solchen Fällen wird ein Block-Design verwendet, bei dem die experimentellen Bedingungen einzeln in Blöcken präsentiert werden. (Sommer 2010, 252). Betrachten wir zum besseren Verständnis wieder ein kleines Beispiel: Carter et al. (Carter et al. 2010, 26–27) geben ein anschauliches Beispiel eines Experiments, bei dem zwei Stimuluskategorien miteinander verglichen werden sollen. Die Kategorie X besteht aus lächelnden Gesichtern, die Kategorie Y aus neutral blickenden Gesichtern. Im ersten Block des Experiments werden der Versuchsperson nur Bilder aus der Kategorie X präsentiert (X1, X2, X3, ...). Ein typischer Block würde 10 bis 60 Sekunden dauern, wobei jeder Stimulus 1 bis 10 Sekunden lang präsentiert werden könnte. Am Ende des ersten Blockes könnte es eine kurze Pause ohne Stimuluspräsentation geben. Danach folgt der zweite Block, der nur Bilder aus Kategorie Y präsentiert (Y1, Y2, Y3, ...) und zeitlich ähnlich aufgebaut ist wie der erste Block. Diese Stimulusblöcke könnten den Versuchspersonen dann abwechselnd weiter präsentiert werden, solange, bis eine ausreichende Menge an Daten generiert ist. Dieses Design wird häufig genutzt, um ein hohes Signal-Rausch-Verhältnis aus den aktivierten Hirnregionen zu bekommen. Denn durch die Addition der hämodynamischen Antworten wird das Signal in den durch die Stimuli aktivierten Hirnregionen stärker über das Rauschen herausragen, als einzelne Antworten allein. Das ist vor allem für solche Experimente von Vorteil, mit denen ermittelt werden soll, welche Hirnregionen bei einem bestimmten Stimulus aktiv werden.

Block-Designs eignen sich weniger für die Erforschung von kurzzeitig auftretenden Änderungen der Hirnaktivitäten, zum Beispiel wenn die Versuchspersonen im Scanner Entscheidungen treffen sollen. Dann ist eine Überlagerung der hämodynamischen Antworten hinderlich bei der Interpretation der Daten. Sommer (Sommer 2010, 253) erklärt, dass speziell für solche Experimente in den 90er Jahren das *Event-related-Design* entwickelt wurde. Da man nicht möchte, dass sich die hämodynamischen Antworten addieren, wird zwischen den Stimuluspräsentationen gewartet, bis die vorherige Antwort abgeklungen ist. Allerdings werden nicht die 25 bis 30 Sekunden bis zum völligen Abklingen abgewartet, sondern lediglich bis der Peak der Antwort vorbei ist, also 10 bis 15 Sekunden. So kann die Höhe des BOLD-Effekts auf einzelne Stimuli gemessen werden. Allerdings mit dem Nachteil, dass das Signal-Rausch-Verhältnis geringer ist als bei Block-Designs.

Mit der Zeit hat sich herausgestellt, dass eine Wartezeit von 10 bis 15 Sekunden zwischen den Stimuli für die Versuchspersonen auf Dauer sehr ermüdend sein kann. Daher kamen Variationen des *Event-related-Designs* auf, die kürzere Interstimulus-Intervalle bis wenigstens 2 Sekunden verwenden. Dabei addieren sich die hämodynamischen Antworten zwar, allerdings nur linear, sodass die einzelnen Peaks noch relativ leicht herauszurechnen sind. Diese *rapid Event-related-Designs* können noch weiter variiert werden, beispielsweise über das sogenannte Jittern, bei dem die Länge der Interstimulus-Intervalle variiert wird. Die jedes Mal veränderte Wartezeit bis zum nächsten Stimulus reduziert Erwartungseffekte, die bei den Versuchspersonen aufkommen. In modernen fMRT-Studien werden die Interstimulus-Intervalle flexibel gestaltet und die Stimuli innerhalb gewisser Grenzen randomisiert präsentiert. (Sommer 2010, 253).

Nachdem das Experiment-Design ausgearbeitet ist, sind auch eventuelle Anforderungen an die Versuchspersonen bekannt. Je nach Studie kann die Rekrutierung geeigneter Versuchspersonen in ausreichender Anzahl ein langwieriger Prozess sein, beispielsweise, wenn ganz bestimmte neurologische Vorerkrankungen gesucht werden. Viele neuroökonomische Studien jedoch benötigen gesunde Menschen ohne mentale Vorerkrankungen, oft alle rechts- oder linkshändig, für eine maximale Vergleichbarkeit. Solche Versuchspersonen finden sich häufig direkt an der Universität oder Forschungseinrichtung, weshalb ein Großteil der Probanden Studenten sind. Die Angaben für eine geeignete Stückzahl pro Studie reichen von circa 10 (Carter et al. 2010, 29) bis circa 25 (Sommer 2010, 255). Im anglo-amerikanischen Raum muss ein *Institutional Review Board (IRB)* über eine Erlaubnis entscheiden, ob ein Experiment mit menschlichen Probanden durchgeführt werden darf. Carter et al. (Carter et al. 2010, 28) erläutern, dass solche IRBs

für gewöhnlich aus ungefähr zehn Personen bestehen, meist Mediziner und Mitarbeiter der Forschungseinrichtung oder des betreffenden Fachbereichs, sowie nicht wissenschaftlich tätige Personen wie Krankenschwestern, Anwälte oder Masterstudierende. Sie sollen darüber entscheiden, ob die vorgelegten Studien ethischen und wissenschaftlichen Vorgaben entsprechen und die Sicherheit der Probanden gewährleistet ist. Die Vorgaben stammen aus dem US-amerikanischen *Health Insurance Portability and Accountability Act* von 1996.

Bereits vor dem Experiment im Scanner treffen sich die Experimentatoren mit den Versuchspersonen, um sie über den Ablauf des Experiments zu informieren und zu erläutern, was von ihnen verlangt wird. Sollten Missverständnisse oder Verständnisschwierigkeiten auftreten, wird so keine teure Scannerzeit verbraucht, um sie zu lösen. Neben der Erklärung ihres Einverständnisses, bei der Studie mitzuwirken und ihre Ergebnisse zu Forschungs- und Veröffentlichungszwecken zur Verfügung zu stellen, müssen die Versuchspersonen auch Angaben über ihre mentale und körperliche Gesundheit, ihren Konsum von Alkohol und Drogen, sowie über metallene Gegenstände oder Flüssigkeiten in und am Körper machen. Das können beispielsweise Herzschrittmacher, künstliche Gelenke oder Schienen sein, aber auch in manchen Tattoofarben befinden sich Metallpartikel. Da fMRT-Scanner mit Magnetfeldern von bis zu 12 Tesla arbeiten, müssen Metalle jeder Art so gut es geht entfernt und vor dem Scannerraum abgelegt werden. Denn magnetisierbare Metalle werden vom Scanner angezogen, vorzugsweise in die Mitte der Röhre. Das kann für die Versuchspersonen gefährlich werden, da der Scanner stark genug ist, auch große Gegenstände wie Gasflaschen anzuziehen, die sich im Raum befinden. Für Gelenke oder Schrittmacher werden zwar meist nicht-magnetisierbare Metalle verwendet. Diese werden vom Scanner auch nicht angezogen, durch das starke Magnetfeld aber erhitzt, sodass Schrittmacher und Gelenke nicht nur verformt und beschädigt werden könnten, sondern auch Verbrennungen drohen.

Das Metall-Problem betrifft nicht nur die Versuchspersonen, sondern auch alle Gegenstände im Scannerraum. Das heißt, dass auch Hilfsmittel, die für das Experiment benötigt werden, nicht aus Metall gefertigt sein dürfen. Sommer berichtet (Sommer 2010, 247), dass sich bereits eine Anzahl an Firmen auf die Herstellung von metallfreien Geräten extra für den Gebrauch in Scannerräumen spezialisiert haben. Diese Produkte beschreibt er allerdings als recht teuer, sodass die Experimentatoren nicht selten auf Einfallsreichtum und handwerkliches Geschick angewiesen sind, um eine Studie durchführen zu können. So hat es sich beispielsweise bewährt, am Kopfende des Scanners in der Wand ein kleines Loch zum Nebenraum anzubringen. In diesem Nebenraum steht dann der Projektor, der die Experimentaufgabe auf einen halbtransparenten Schirm projiziert, der

am Kopfende der Scannerröhre angebracht ist. Über dem Kopf der Versuchsperson wird dann ein Spiegel leicht schräg aufgehängt, mit dem sie sehen kann, was auf dem Schirm hinter ihrem Kopf zu sehen ist. (Sommer 2010, 244). Das ist etwas umständlicher als eine Spezialbrille für die Probanden, aber sicherlich kostengünstiger. Damit die Versuchspersonen auf den Stimulus reagieren, also eine Entscheidung treffen können, bekommen sie eine Art Fernbedienung mit Knöpfen (je weniger desto besser) in die Hand. Durch Drücken beispielsweise des rechten oder linken Knopfes können sie ihre Entscheidungen treffen. Carter et al. (Carter et al. 2010, 30) beschreiben auch, dass ihre Probanden immer in die andere Hand einen quetschbaren Ball bekommen. Wenn sie sich unwohl fühlen, kann der Ball gedrückt werden und das Experiment wird sofort abgebrochen. Denn nicht jede Versuchsperson kann sich lange im Scanner aufhalten. Zum einen ist es recht eng, zum anderen recht laut, weil der Scanner während des Betriebs starke Geräusche macht. Daher könnten es die Experimentatoren auch nicht immer hören, wenn eine Versuchsperson etwas äußert, sei es nun eine Antwort auf eine Experimentaufgabe oder den Wunsch zum Abbruch der Session.

Zur Enge im Scanner trägt auch bei, dass Kopfbewegungen der Versuchspersonen so gut wie möglich verhindert werden müssen. Denn zu Beginn des Experiments wird ein Bild des Gehirns der Person aufgenommen, auf das die funktionellen Daten aufgebracht werden. Verrutscht die Position des Kopfes relativ zu diesem Referenzbild, könnte das die Positionen der Aktivierungen verfälschen. Daher wird der Kopf von Versuchspersonen so gut wie möglich mit Kissen und Polstern fixiert, manchmal sogar mit einer Beißschiene oder Kissen, die um den Kopf herum aufgeblasen werden.

Die Experimentaufgabe, die sorgfältig ausgearbeitete Präsentation der Stimuli, wird in eine eigene Software programmiert, die mit dem Scannercomputer synchronisiert werden muss, damit der Ablauf des funktionellen Scans mit dem Ablauf der Präsentation, und damit der wertvollen Arbeit der Versuchspersonen übereinstimmt. Sonst können die verschiedenen Aktivierungsmuster später nicht bestimmten Zeitpunkten in der Stimuluspräsentation zugeordnet werden. (Sommer 2010, 247). Sind die Versuchspersonen ausgestattet und der Scanner sowie die Computer bereit, kann das Scannen beginnen.

An dieser Stelle kommt unsere Beispiel-Studie aus der Neuroökonomie hinzu, eine Studie von Jamil Zaki und Jason Mitchell, Psychologen an der Universität Harvard (Zaki et al. 2011). An dieser Studie werden wir den Versuchsdurchlauf und die Analyse in einzelnen Schritten nachvollziehen. Diese Studie ist nicht nur deshalb ein gutes Beispiel, weil sie relativ moderne Experimentstandards mit fMRT-Geräten veranschaulicht, sondern auch, weil sie direkt aus dem Bereich der neuroökonomischen Anwendung kommt und inhaltlich ein gutes

Beispiel dafür ist, wie die Zusammenarbeit von Verhaltensökonominnen und kognitiven Neurowissenschaftlern konkret aussehen könnte, wie wir weiter unten sehen werden,

Was hat es mit dieser Studie von Zaki und Mitchell also auf sich? Sie untersucht die neuronalen Implementierungen gerechter Entscheidungen und ist ein schönes Beispiel für neuroökonomische Forschung. Die Autoren (Zaki et al. 2011, 19761) bemerken, dass ökonomische Standardmodelle von egoistischen Menschen ausgehen, jedoch in jüngeren Studien regelmäßig faires und altruistisches Verhalten beobachtet wird. Laut den Autoren wird solches Verhalten von den Standardmodellen mit einem äußeren Druck zu erklären versucht, die eigene Reputation zu schützen und Vergeltungshandlungen zu vermeiden. Ihrer Ansicht nach widersprechen diesem Erklärungsversuch Beobachtungen, dass Versuchspersonen beim Diktatorspiel Geld an ihre Mitspieler abgeben, ohne Sanktionen oder Schädigungen ihres Rufs befürchten zu müssen. Sie erläutern eine alternative Erklärung dieses Verhaltens, nach der soziale Prinzipien wie Fairness, Austausch und Kooperation einen intrinsischen Wert besitzen, sodass Individuen zwar ihren eigenen Vorteil wertschätzen, aber auch eine Aversion gegen Ungerechtigkeiten und Egoismus empfinden. Weiter berichten sie, dass in jüngerer Zeit eine wachsende Unterstützung dieser Hypothese aus bildgebenden Studien an Menschen kommt: Es wurde verlässlich gezeigt, dass Aktivitäten in Regionen wie dem ventralen Striatum und dem orbitofrontalen Kortex ansteigen, linear zum subjektiven Wert einer Reihe von Belohnungen. Diese Hirnregionen sind laut jüngerer Studien auch dann aktiv, wenn die Versuchsperson soziales Verhalten, wie gerechte Ressourcenverteilung, beobachten. Zaki und Mitchell sehen aber bei diesen Studien den Nachteil, dass sie lediglich passives Beobachten untersuchten, die Empfindungen von Entscheidungsträgern selbst jedoch nicht bekannt sind. Diese möchten sie in ihrer Studie untersuchen, der die Idee zugrunde liegt, dass, wenn Gerechtigkeit von Entscheidungsträgern als belohnend empfunden wird, ein Festhalten an diesem Prinzip, unabhängig vom eigenen Gewinn, Hirnregionen aktivieren sollte, die mit Wertschätzungen und Belohnungen zusammenhängen. Sollten gerechte Entscheidungen aber von der Unterdrückung vorherrschender egoistischer Impulse herrühren, sollten sie von Aktivitäten in Hirnregionen begleitet sein, die kognitive Kontrolle ausüben.

Für ihre Studie rekrutierten Zaki und Mitchell offenbar eine Reihe Studenten, denn das Durchschnittsalter der neunzehn Versuchspersonen (zwölf männlich, sieben weiblich) betrug 23,2 Jahre. Alle waren rechtshändig und hatten keine Historie psychiatrischer oder neurologischer Erkrankungen. Für die Teilnahme an der Studie wurde ihnen eine monetäre Kompensation geboten, deren Höhe

sich aus einem Grundbetrag und einem zusätzlichen Geldbetrag zusammensetzte, der von ihren Entscheidungen im Experiment abhing.

Als experimentelles Paradigma nutzten sie ein Diktatorspiel, das sie auf *Event-related-Design* mit faktoriellem Design modifiziert hatten. Vor dem Scannen wurde immer eine Versuchsperson mit einem Mitarbeiter des Experiments zusammengebracht, der der Versuchsperson als zweiter Versuchsteilnehmer vorgestellt wurde. Die Experimentatoren erklärten beiden, in der Studie ginge es um die Frage, wie Individuen Eindrücke über reale andere Personen formen, denen sie tatsächlich begegnet sind, im Gegensatz zu fiktionalen anderen Personen. Eine der beiden Versuchspersonen sollte in den Scanner liegen, während die andere in einem anderen Raum davon unabhängige Aufgaben erledigen sollte. Eine ökonomische Entscheidungsaufgabe wurde nicht erwähnt. Dann wurde aus den Teams der Versuchspersonen in einem Verfahren, das der Versuchsperson zufällig erschien, immer die tatsächliche Versuchsperson für die Scanneraufgabe ausgewählt. Diese Versuchsperson nahm im Diktatorspiel die Rolle des Diktators ein, die andere Person die Rolle des Empfängers. Im Scanner wurden die Versuchspersonen darüber informiert, dass sie eine Reihe von Entscheidungen treffen sollten. Sie bekamen die Wahl, eine bestimmte Summe Geld entweder sich selbst oder dem Empfänger zukommen zu lassen. Fünf ihrer Entscheidungen würden hinterher zufällig ausgewählt und tatsächlich umgesetzt. Um zu verhindern, dass die Versuchspersonen mit dem Gedanken an ihren Ruf oder moralische Reputation handelten, wurde ihnen versichert, dass der Empfänger nicht erfahren würde, dass die Versuchsperson im Scanner das Diktatorspiel gespielt hatte. Der zusätzliche Geldbetrag, den der Empfänger eventuell aus dem Spiel erhalten würde, werde ihm einfach zusammen mit der Grundkompensation zugesandt.

Die Versuchspersonen durchliefen drei Sessions im Scanner mit insgesamt 210 Runden des Diktatorspiels, das folgendermaßen ablief: Nach den Instruktionen wurden die Versuchspersonen im Scanner platziert. Das Experimentdesign sah zunächst ein gejittertes Interstimulus-Intervall vor, dessen Länge von 1 bis 6 Sekunden reichen konnte. Danach wurde ihnen der Stimulus in zwei Wahlmöglichkeiten präsentiert. Auf dem Bildschirm erschienen am linken und am rechten Rand je ein Foto, das eine zeigte die Versuchsperson, das andere den Empfänger. Welches Foto rechts oder links zu sehen war, variierte zufällig bei jedem Trial. Über den Fotos stand jeweils „\$XX.XX for“, mit variierenden Geldbeträgen. Dieses Bild blieb für ein wieder gejittertes Zeitintervall von 1 bis 5 Sekunden zu sehen. Dann kam die Aufforderung zu entscheiden über das Wort „Decide“, das zusätzlich auf dem Bildschirm erschien. Von da an hatten die Versuchspersonen 2 Sekunden Zeit, über die Knöpfe ihrer Fernbedienung eine Wahl zu treffen, wer den Geldbetrag bekommen sollte. Das Foto der gewählten Person erschien

dann für den Rest der Runde rot umrahmt. Danach begann mit einem erneuten gejjitterten Interstimulus-Intervall eine neue Runde.

Die Geldbeträge, die die Diktatoren vergeben konnten, variierten, lagen aber immer zwischen \$0,00 und \$3,00. Die Relationen zwischen den Beträgen hielten sich dabei immer an eine vorgeschriebene Reihe von Möglichkeiten (3:1, 2:1, 3:2, 4:3, 5:4, 1:1), wobei der mögliche Betrag für den Diktator zuerst gewählt wurde. Anhand einer zufällig gewählten Relation wurde dann der mögliche Betrag für den Empfänger errechnet.

Beispielsweise konnte ein Stimulus so aussehen, dass über dem Foto der Versuchsperson „\$1.00 for“ stand und über dem Foto des Empfängers „\$1.33 for“. Die Diktatoren konnten dann zwischen links und rechts entscheiden, also ob sie selbst \$1 bekommen sollten und der jeweilige Empfänger dann leer ausginge, oder ob der Empfänger \$1,33 erhalten sollte und sie selbst keine Auszahlung bekämen. Auf diese Weise konnten vier Faktoren implementiert werden, die danach unterschieden wurden, ob der höhere Betrag für die Diktatoren oder die Empfänger möglich gewesen wäre. Stand der höhere Betrag bei den Empfängern, wäre es nach Zaki und Mitchell von den Diktatoren gerecht gewesen, den Empfängern diesen Betrag auch zukommen zu lassen (*generous/equitable*), während eine Entscheidung, trotzdem den niedrigeren Betrag ihnen selbst zu gewähren, als eigennützig eingestuft worden wäre (*self-serving/inequitable*). Stand dagegen der höhere Betrag auf der Seite der Diktatoren, wäre eine Allokation bei den Empfängern eine generöse, aber ungerechte Entscheidung gewesen (*generous/inequitable*), während eine Entscheidung zugunsten der Diktatoren selbst eine eigennützige, aber gerechte gewesen wäre (*self-serving/equitable*). In dieser Studie war eine Entscheidung also genau dann gerecht, wenn sie demjenigen Spielpartner das Geld zusprach, der mehr hatte gewinnen können, ganz gleich, ob das die Diktatoren selbst oder die Empfänger waren.

Zudem wurden noch eine Reihe *pure self* und *pure other* Wahlmöglichkeiten eingebracht. Sie wurden als reine Belohnung für die Diktatoren (*pure self*) beziehungsweise die Empfänger (*pure other*) eingebaut (für die je anderen wurden dabei \$0,00 angeboten), ohne das Prozedere des Experiments zu verletzen. Außerdem wurden Situationen eingefügt, in denen die Diktatoren zwischen \$0,00 für sich und \$0,00 für die Empfänger entscheiden konnten, sodass, egal welche Wahl getroffen wurde, niemandem ein Geldbetrag zugesprochen werden konnte. Auch diese Ergebnisse flossen in die Datenanalyse ein. Der Hintergrund dafür könnte sein, dass die Experimentatoren von Zeit zu Zeit ein sogenanntes Null-Event einfügen wollten. Während eines Null-Events soll das BOLD-Signal die Möglichkeit bekommen, wieder auf die Baseline abzusinken.

Sie sind Interstimulus-Intervalle, die oft die Länge eines normalen Events besitzen, nur ohne einen Stimulus zu präsentieren. Die Wahl von \$0,00 für beide Personen könnte also eine Möglichkeit sein, das durch das *rapid Event-related-Design* stets angehobene BOLD-Signal zwischendurch absinken zu lassen, ohne die Gleichförmigkeit des Experimentablaufs zu unterbrechen.

An diesem Beispiel lässt sich gut beobachten, wie ein bekanntes ökonomisches Spiel in ein im Scanner durchführbares Spiel übersetzt wird. In der klassischen ökonomischen Variante des Diktatorspiels erhält der Diktator einen bestimmten Betrag; wie viel sie davon an den Empfänger abgeben möchte, kann sie selbst wählen. In der Version des Spiels von Zaki und Mitchell können die Diktatoren die Höhe des abzugebenden Betrags nicht selbst regulieren, das wäre zu kompliziert für eine Entscheidung im Scanner. Dazu müsste schon das Antwortinstrument (hier die Fernbedienung) mehr als zwei einfache Knöpfe besitzen, die rechts oder links, ja oder nein, codieren. Noch dazu wären die Prozesse im Gehirn dabei viel komplexer und könnten so nur schwer modelliert werden. In der Scannerversion des Spiels gibt es für die Diktatoren nur die Überlegung, ob sie die eine oder die andere Möglichkeit besser finden (also welche der beiden Personen das für sie vorgesehene Geld erhalten soll), während sie im klassischen Spiel theoretisch sehr viele Möglichkeiten hätten, das Geld aufzuteilen. Dafür wäre es schwieriger, zu modellieren, welche Prozesse im Gehirn vor sich gehen, ob etwa für jede der potenziellen Verteilungsvarianten ein eigener Wert berechnet und mit allen anderen verglichen wird. Die Prozesse, die sich während des Experiments im Gehirn abspielen, sollten so unkompliziert wie möglich gehalten werden. Wie wichtig die Möglichkeit einer relativ einfachen Modellierung ist, werden wir weiter unten bei der Datenanalyse mit Regressoren sehen.

Die Versuchspersonen liegen also im Scanner und treffen ihre Entscheidungen. Nach einer Weile wird eine Pause eingelegt, die Versuchsperson trinkt vielleicht einen Kaffee und geht ein paar Schritte im Flur, dann geht es weiter mit der zweiten Session, nach einer weiteren Pause mit einer dritten. Dieses Vorgehen wiederholt sich für alle Versuchspersonen, sofern keine außerplanmäßigen Vorkommnisse eintreten, in unserem Beispiel mit 15 Personen.

Danach dürfen die Versuchspersonen nach Hause gehen, und die Datenanalyse beginnt. Dazu müssen die Daten vorverarbeitet werden, um interpersonelle Vergleichbarkeit zu erreichen. Gute Veröffentlichungen geben nicht nur an, welche Art von Scannergerät verwendet wurde, sondern auch, welche Softwares und Modelle zur Datenverarbeitung. In unserem Beispielfall war der Scanner ein 3-Tesla Siemens Trio-Scanner, die Stimuluspräsentation geschah über MATLAB und die Datenverarbeitung wurde mithilfe der *SPM Statistical Parametric Mapping Image Analysis Software* durchgeführt.

Wie bereits erwähnt, wird gewöhnlich zu Beginn der ersten Session im Scanner ein nicht-funktionelles MRT-Bild des Gehirns aufgenommen, auf das die Aktivierungsdaten aufgebracht werden. Zaki und Mitchell entschieden sich für T1-gewichtete Aufnahmen ihrer Versuchspersonen. Damit die Aktivierungen an der korrekten Stelle dargestellt werden, müssen die Daten zunächst einer Aquisitionszeit- und Bewegungskorrektur unterworfen werden. Die Aquisitionszeitkorrektur wird durchgeführt, weil die Software des Scanners das Gehirn in mehrere virtuelle Schichten aufteilt, deren Stärke vor dem Scannen festgelegt wird (in unserem Beispiel 5 Millimeter). Die Schichten werden nacheinander gescannt, über die sogenannte Repetitionszeit hinweg. In unserem Beispiel liegt diese bei 2 Sekunden. Das bedeutet, dass die letzte Gehirnschicht bei jedem Scan 2 Sekunden nach der ersten Schicht gescannt wird. In diesem Zeitraum können sich die Aktivierungen verändern, was vor allem klar wird, wenn man die Zeitintervalle der Stimuluspräsentation betrachtet. Daher wird der zeitliche Unterschied vom Computer herausgerechnet, sodass für die statistische Analyse angenommen werden kann, dass alle Bilder eines Volumens (3D-Bild des gesamten Gehirns, das während einer Repetitionszeit aufgenommen wurde) gleichzeitig aufgenommen wurden. Diese Korrektur wird nur bei Experimenten mit event-related Designs angewandt, da hier der möglichst exakte Zeitpunkt eines BOLD-Signals eine Rolle spielt. (Sommer 2010, 256).

Die Bewegungskorrektur ist notwendig, weil sich die Köpfe der Versuchspersonen im Scanner trotz aller Vorsicht während einer Session bewegen. Da die Voxel im Scanner ein statisches Raster bilden, können nach einer Kopfbewegung ganz andere neuronale Verbände in einem Voxel zu finden sein als vorher. Bei der Bewegungskorrektur wird also jedes einzelne Volumen gedreht und gekippt, bis es wieder exakt über das Referenzvolumen (in unserem Fall das T1-gewichtete Bild vom Anfang des Experiments) passt. Für jedes Volumen werden die vorgenommenen Bewegungsparameter zur späteren Ansicht gespeichert. (Sommer 2010, 256).

Ein weiterer wichtiger Schritt der Datenvorverarbeitung, den auch Zaki und Mitchell vornahmen, sind die Normalisierung und Glättung der Daten. Da sich jedes Gehirn in seiner anatomischen Struktur von anderen Gehirnen unterscheidet, werden die Daten aller Versuchspersonen auf eine Art Mustergehirn-Schablone aufgebracht. Dazu werden die strukturellen Daten des Gehirns durch Stauchen und Strecken an die Strukturen des Mustergehirns angepasst. Für solche Normalgehirne gibt es mehrere Angebote, eines der bekanntesten kommt vom *Montreal Neurological Institute* (MNI), der *MNI Space*. (Carter et al. 2010, 34; Sommer 2010, 256). Zaki und Mitchell nutzten von diesem Institut das *International*

*Consortium of Brain Mapping ICBM 152 brain template.*⁴ Als letzter Schritt folgt dann noch die räumliche Glättung der Daten über einen Gaußschen Filter. Dabei werden die gemessenen Aktivierungswerte für die einzelnen Voxel mit ihren Nachbarvoxeln verglichen. Die Intensität eines Voxels wird auf die umgebenden Voxel ausgebreitet, sodass einzelne Ausschläge geglättet werden. Dazu legen die Forscher die Weite des Filters fest. Sie bestimmt die Weite der Wirkung des Glättungseffekts. Ein enger Filter verteilt die Daten nur über ein paar Voxel, während ein weiter Filter sie über viele Voxel verteilt. Die Filtergröße wird üblicherweise in Millimeter beim halben Maximum der Gaußschen Kurve (*Full-Width-Half-Maximum*, FWHM) angegeben und sollte mindestens die zwei- bis dreifache Voxelgröße⁵ betragen. In unserem Fallbeispiel verwendeten die Autoren einen Filter von 6 mm FWHM. Die räumliche Glättung soll zum Einen noch immer bestehende Unterschiede zwischen den Probanden ausgleichen, zum Anderen, das ist wichtig, soll sie das Signal-Rausch-Verhältnis verbessern, denn eine durch den Stimulus bedingte Aktivierung erreicht meist zwei bis drei benachbarte Voxel, sodass es bei der Beurteilung der Aktivierung eines Voxels hilfreich ist, die Aktivierung der Nachbarvoxel miteinzubeziehen. Dadurch wird auch die Rate von falsch positiven Aktivierungsanzeigen verringert. Ohne Glättung können einige einzelne Peaks von Aktivierungen den Grenzwert für signifikante Peaks überschreiten, obwohl sie nicht durch den Experimentstimulus ausgelöst wurden, sondern aus dem Rauschen stammen, und daher für die Analyse nicht berücksichtigt werden sollten. Bei der Glättung fallen diese einzelnen Peaks oft durch, weil ihre Nachbarvoxel keine oder zu wenig Aktivierung aufweisen. So steigt die Wahrscheinlichkeit, dass es sich bei den Aktivierungen, die nach der Glättung noch signifikant scheinen, auch tatsächlich um für die Analyse wertvolle Aktivierungen handelt. (Huettel et al. 2014, 316–317).

Nach der Datenvorverarbeitung kann die eigentliche Datenanalyse beginnen. Zuvor noch ein Wort zu den behavioralen Ergebnissen der Studie von Zaki und Mitchell: Von den Trials, in denen den Versuchspersonen mehr Geld angeboten wurde als den Empfängern, wiesen die Versuchspersonen das Geld in 83,2% der Fälle sich selbst zu (*self-serving/equitable*). Bei den Trials, in denen den Empfängern mehr Geld angeboten wurde, gaben es die Versuchspersonen mit ungefähr der gleichen Wahrscheinlichkeit an die Empfänger weiter, wie sie es sich selbst

⁴ Die Erwähnung technischer Details dieser Studie wird uns im folgenden Kapitel 6 noch nützlich sein.

⁵ Für Voxel gibt es keine festgelegten Größen. Welche Größe ein Voxel haben soll, wird von den Experimentatoren bestimmt. Sie sollten weder zu groß sein, da sie dann zu viele Neuronen beinhalten könnten, noch zu klein, da dann zu viele Voxel pro Gehirnsan anfallen, die bearbeitet werden müssen. Beides könnte die Messungen ungenau werden lassen.

zuteilt (*generous/equitable* 55,2%; *self-serving/inequitable* 44,6%). Entscheidungen, bei denen den Versuchspersonen mehr Geld angeboten wurde, diese es aber an die Empfänger weitergaben (*generous/inequitable*), kamen so selten vor, dass sie in der Analyse nicht berücksichtigt werden konnten. Insgesamt gaben die Versuchspersonen 22,2% des insgesamt verfügbaren Geldes an die Empfänger weiter, was, wie die Autoren sagen, mit Ergebnissen früherer Studien mit Diktatorspielen im Einklang steht. (Zaki et al. 2011, 19762).

Zurück zur Datenanalyse: Für die Analyse der Daten nutzten Zaki und Mitchell eine Regressionsanalyse mit dem Allgemeinen Linearen Modell (ALM). Diese Art der Analyse wird für die meisten fMRT-Studien verwendet und ist in vielen Softwarepaketen für fMRT-Analysen enthalten. Für eine solche modellbasierte Analyse versuchen die Experimentatoren, die aufgezeichneten Aktivierungsdaten so gut wie möglich mit Regressoren zu erklären. Regressoren sind hypothetische Verläufe von BOLD-Antworten in einem Voxel über einen bestimmten Zeitraum. Sie korrespondieren mit verschiedenen Prozessen, oder Teilen davon, die sich hypothetisch im Gehirn abspielen (beispielsweise phonologischer Abruf). Das ALM kommt üblicherweise in Matrizenform und bringt die Regressoren mit den Daten in Verbindung:

$$y = G \times \beta + \varepsilon,$$

wobei y die beobachteten Daten sind, G ist die Design-Matrix, β die Parameter Matrix und ε die Error Matrix. Die Error-Matrix enthält die Fehler für jedes Voxel, die durch Rauschen in den Daten oder Fehler bei den Messungen entstehen. Die Daten y sind bereits bekannt, sie wurden gerade gemessen. Die Idee beim ALM ist, dass die Experimentatoren die Regressoren so auswählen, dass sie möglichst viele der aufgetretenen Daten möglichst genau vorhersagen. Dazu müssen die Experimentatoren eine Vorstellung davon haben, welche Aktivierungen in einem Voxel über den Zeitverlauf auftreten. Die Gesamtheit der Regressoren wird in der Design-Matrix G festgehalten. Zusätzlich werden die Regressoren gewichtet. Je höher die Gewichtung für einen Regressor, desto stärker korreliert er mit den beobachteten Daten. Diese Gewichtung geschieht über die Parametergewichte β , die in der Parameter-Matrix festgehalten sind. (Huettel et al. 2014, 372–374).

Die Regressoren sind hypothetische Erklärungsversuche, die von den Experimentatoren ausgewählt werden, ebenso wie ihre Gewichtungen. Zusammen können sie die Daten gut erklären, oder schlecht. In letzterem Fall ist die Fehlervarianz des Modells sehr hoch, sie sollte aber so gering wie möglich gehalten werden. Denn, wie Sommer (Sommer 2010, 259) erklärt, werden die Aktivierungen später einzeln danach beurteilt, ob sie für die Forschungshypothese

relevant oder zufällig entstanden sind. Dabei wird nach der Höhe der Fehlervarianz geschaut, die die Gewichtung des Regressors für diese Aktivierung relativiert. Je höher die Fehlervarianz, desto höher die Wahrscheinlichkeit für eine zufällige Aktivierung.

Zaki und Mitchell (Zaki et al. 2011, 19765) modellierten mit ihrem ALM die Entscheidungsphase ihres Experiments. Dazu wählten sie drei Regressoren, die den drei interessierenden Entscheidungsbedingungen entsprachen: *self-serving/inequitable*, *self-serving/equitable* und *generous/equitable*. Die vierte Bedingung (*generous/inequitable*) war, wie gesagt, zu selten aufgetreten, um modelliert werden zu können. Die *Whole-Brain-Analyse* verglich die Bedingungen in der Weise [*self-serving/equitable* + *generous/equitable*] > *self-serving/inequitable* und gegensätzlich dazu *self-serving/inequitable* > [*self-serving/equitable* + *generous/equitable*]. Dadurch ließen sich Häufungen von Gehirnaktivitäten herausfiltern, die zwischen gerechten Entscheidungen (beider Art) und ungerechten Entscheidungen unterschieden. Die Höhe des Geldbetrags, der bei jedem Trial gewonnen werden konnte, hatten Zaki und Mitchell als Faktor ohne Interesse in ihr Modell eingearbeitet. Daher gehen sie davon aus, dass die Analyse die Auswirkungen bestimmter Entscheidungstypen aufzeigt, ohne vom möglichen Geldgewinn beeinflusst zu sein. Diese Analyse zeigte Aktivierungen des orbitofrontalen Kortex (OFC) im Zusammenhang mit gerechten, im Gegensatz zu ungerechten, Entscheidungen auf. Der OFC reagierte dabei sowohl auf egoistische, als auch auf großzügige Entscheidungen, solange sie gerecht (*equitable*) waren. Zaki und Mitchell leiteten daraus eine erste Hypothese ab (Zaki et al. 2011, 19762):

„Thus, OFC activity here cannot have reflected (i) the presence or absence of personal gain or (ii) the ‚warm glow‘ of acting generously per se, because generous/equitable and self-serving/equitable trials differed along both of these dimensions but engaged this region similarly. Remarkably, the lowest OFC response was observed when participants chose to allocate money inequitably to themselves, even though these choices resulted in real financial gain for the participant, suggesting that a motivation to maximize social outcomes may ‚crowd out‘ the value associated with personal gains.“

Diese Aktivierungen wollten Zaki und Mitchell noch weiter untersuchen. Trat die gerechtigkeitsbezogene Aktivität in einer Region des OFC auf, die generell für Werte aktiv wird? Für diese Frage analysierten sie die Antworten auf die *Pure-Self-Trials*. Diese Trials verschafften den Versuchspersonen reinen Gewinn, ohne Kosten für die Empfänger und ohne Möglichkeit für die Versuchspersonen, großzügig zu entscheiden. Für diese Analyse wurde wieder das ALM

verwendet, diesmal mit einem Regressor, der die Höhe des Geldbetrages repräsentierte, den die Versuchspersonen in den einzelnen *Pure-Self-Trials* erhalten hatten. Die Analyse erbrachte, dass eine ähnliche Region im OFC wie oben aktiv wurde, wobei die Aktivierungen mit der Höhe der Geldbeträge korrelierte. Diese dabei identifizierte *Region of Interest* im OFC wurde daraufhin nocheinmal nach unterschiedlichen Aktivierungen für die unterschiedlichen Entscheidungsmöglichkeiten untersucht. Die Aktivierungen waren stärker für gerechte gegenüber ungerechten Entscheidungen, wobei sie sich wiederum nicht zwischen eigennützig gerechten und großzügigen gerechten Entscheidungen unterschieden. Für Zaki und Mitchell (Zaki et al. 2011, 19762) unterstützt das die Schlussfolgerung, dass Entscheidungsträger Gerechtigkeit einen intrinsischen Wert zuschreiben.

Um ihren Punkt weiter zu untermauern, unternehmen die Autoren eine Art Gegenprobe (Zaki et al. 2011, 19762, 19766). Ihre Idee ist, dass wenn gerechte Entscheidungen mit einem subjektiven Wert verbunden sind, ungerechte Entscheidungen ein Gefühl von subjektiver Disutilität hervorrufen sollten. Eine Gegenüberstellung der Gehirnaktivitäten für gerechte und ungerechte Entscheidungen erbrachte für ungerechte Entscheidungen erhöhte Aktivitäten in der vorderen Insula, die gemeinhin mit aversiven Gefühlen wie Schmerz oder Abscheu in Verbindung gebracht wird. Zudem zeigte sich, dass diejenigen Versuchspersonen mit stärkerer Insulaaktivierung bei ungerechten Entscheidungen weniger ungerechte Entscheidungen fällten. Da diese Personen somit auch weniger Daten zu ungerechten Entscheidungen bereitstellten, überprüften Zaki und Mitchell diesen Befund anhand einer weiteren ALM-Analyse eigens für die vordere Insula, bei der sie für jede Versuchsperson genau zehn ungerechte Entscheidungen berücksichtigten, um die Vergleichbarkeit der Daten zu erhöhen. Diese Analyse kam zu dem gleichen Ergebnis wie die vorherige *Whole-Brain*-Analyse. Damit war die Datenanalyse für diese Studie abgeschlossen, es folgt die Interpretation der Ergebnisse.

Für ihre Studie kamen Zaki und Mitchell (Zaki et al. 2011, 19763) zu dem Schluss: „[...] charitable donation engages reward-related brain regions.“ Zudem sehen sie in ihrer Analyse der vorderen Insula Belege, die nahelegen, dass nicht nur die Beobachtung, sondern auch das eigene ungerechte Handeln als aversiv empfunden wird.

Für unsere Fragestellung ist dieses Experiment interessant, weil es die ursprünglich ökonomische Fragestellung nicht lediglich aus einer neurowissenschaftlichen Sicht beleuchtet, es fügt vielmehr etwas Neues hinzu. Zaki und Mitchell bearbeiteten die Frage, ob Versuchspersonen auch ohne (sozialen) Druck von außen beim Diktatorspiel Geld abgeben. Diese Frage glauben sie beantworten zu können, ist für uns aber nur von untergeordneter Bedeutung. In unserem

Kontext ist es vielmehr interessant, dass die Studie eine Erweiterung der Möglichkeiten darstellt, die den Ökonomen in ihren eigenen Laboren nicht zur Verfügung gestanden hätte. Mit behavioralen Experimenten ließe sich zwar der Umstand aufdecken, dass Versuchspersonen im Diktatorspiel ohne Zwang von außen Geld an die Empfänger abgeben. Aus welchen Gründen das aber geschieht, lässt sich nur schwer feststellen. Immerhin kann man im Experiment die Angst vor Reputationsschädigungen und Vergeltungen ausschließen. Aber die Frage, ob die Diktatoren aus einem *Warm-Glow*-Empfinden heraus Geld abgeben, lässt sich kaum beantworten. Hier konnte die Studie von Zaki und Mitchell immerhin einen Anfangspunkt liefern. Mit dem fMRT und einem geschickten Experimentdesign hatten sie die Möglichkeit, Erklärungsversuche für die Gründe nicht nur zu liefern, sondern auch ein Stück weit zu untermauern. Natürlich ist ihre Studie nicht der Wahrheit letzter Schluss zu dieser Frage, aber sie erweitern die möglichen Fragestellungen, und damit Erkenntnismöglichkeiten, zu ökonomischen Problemen.

Hier liegt der Beitrag der kognitiven Neurowissenschaften zur Neuroökonomie verborgen. Zum einen sind es die Daten, behavioral wie neurologisch, zum Verhalten von Entscheidungsträgern in ökonomischen Situationen. An sich können diese schon den Grundstock an Erkenntnissen erweitern. Der wahre Wert liegt dabei aber in den Möglichkeiten der neurologischen Instrumente und Fragestellungen, die über die ökonomischen Möglichkeiten hinausgehen, wie man an der Studie von Zaki und Mitchell sehr schön sehen kann, und was wir im folgenden Abschnitt 5.3 genauer sehen werden. Mit behavioralen Methoden kann man den Leuten bekanntermaßen nur vor den Kopf schauen. Inwiefern die bunten Bilder aus einem fMRT dazu geeignet sind, menschliches Verhalten nicht nur zu erklären, sondern global zu modellieren und schlussendlich auch vorherzusagen, ist natürlich eine andere Frage, der wir uns in Kapitel 6 stellen müssen. Werden die Möglichkeiten der kognitiven Neurowissenschaften aber nicht in prophetischer, sondern wissenschaftlicher Weise eingesetzt, könnten die kognitiven Neurowissenschaften mit ihren Studien einen wichtigen und wertvollen Beitrag zur Neuroökonomie leisten, wie wir nun genauer betrachten werden.

5.3 Wie nutzen Verhaltensökonomie und kognitive Neurowissenschaften die Vorteile des jeweils anderen?

In den Kapiteln 3 und 4 haben wir gesehen, welche Ideen und Entwicklungen hinter der Verhaltensökonomie beziehungsweise den kognitiven Neurowissenschaften stecken und wonach sie suchen. Kapitel 2 ließ erkennen, welche Hoffnungen die Anhänger der neuroökonomischen Idee in die Neuroökonomie stecken. Das vorliegende Kapitel zeigt nun auf, wie die Arbeit aussehen kann, die die beiden Wissenschaften in die Neuroökonomie einbringen. Nun wollen wir sehen, welchen Nutzen sie aus der Neuroökonomie ziehen können, was sie voneinander lernen und anwenden können.

Diese Frage zu beantworten, erfordert einen breiten Blick auf die Neuroökonomie. Je nachdem, wen man fragt, erhält man unterschiedliche Antworten, worin der tatsächliche Nutzen der Neuroökonomie liegt. Da sind auf der einen Seite diejenigen Ökonomen und Neurowissenschaftler, die die Neuroökonomie runderaus ablehnen, aus verschiedenen Gründen, die wir in Kapitel 6 betrachten werden. Auf der anderen Seite sieht eine Gruppe von Forschern großen Nutzen und vielfältige Möglichkeiten in der Neuroökonomie. Bekannte Vertreter sind etwa Colin Camerer, Paul Glimcher oder Ernst Fehr. In diesem Kapitel wird die Sichtweise dieser Anhänger und Anwender der Neuroökonomie erläutert. Daneben gibt es aber auch eine Vielzahl an Verhaltensökonomien und kognitiven Neurowissenschaftlern, die sich weder zu den glühenden Verfechtern, noch zu den entschiedenen Gegnern der Neuroökonomie zählen würden. Sie lehnen die Neuroökonomie nicht grundsätzlich ab, sondern sehen sie eher als möglicherweise nützliche Zusammenarbeit. Diese Gruppe gilt es für die Anhänger der Neuroökonomie zu überzeugen. Das war ihnen bereits vor zehn Jahren klar, als die Neuroökonomie gerade an Fahrt aufnahm. Damals entwickelte sich eine hitzige Debatte darüber, welchen Nutzen die Neuroökonomie für ihre Elternwissenschaften haben könnte. Inzwischen ist die Diskussion etwas abgekühlt, die Frage besteht aber noch immer und die heutigen Ansichten sind natürlich von den Versprechen geprägt, die damals gemacht wurden. Manche dieser Versprechen über den zukünftigen Nutzen der Neuroökonomie klingen realistisch und wecken Interesse, einige klingen aber auch sehr optimistisch, um nicht zu sagen, recht hochgegriffen und unrealistisch. Bei manchen Versprechen wird auch nicht klar, ob sie aus grundsätzlichen, strukturellen, wissenschaftstheoretischen Gründen überhaupt erfüllbar sind. Im folgenden Kapitel 6 werden wir uns eingehender mit den Versprechen und ihrem Echo in der wissenschaftlichen Gemeinschaft

befassen. Daher genügt uns an dieser Stelle ein kurzer Blick auf die Versprechen vom Nutzen der Neuroökonomie, bevor wir sehen, was davon übrig ist.

Möglicher Nutzen aus der Neuroökonomie geht grundsätzlich in eine von zwei Richtungen: Richtung Ökonomie oder Richtung Neurowissenschaften. Es gibt keinen Konsens darüber, in welche Richtung mehr Nutzen fließen kann oder wird. Nach Ansicht vieler, vor allem neutral bis skeptisch eingestellter Ökonomen, können die Neurowissenschaften mehr von der Expertise der Ökonomen profitieren, wenn überhaupt, während Neurowissenschaftler den höheren potenziellen Wissenstransfer in Richtung Ökonomie sehen. Die große Debatte aus den Anfängen der Neuroökonomie, die wir hier betrachten, drehte sich vor allem um die Frage des Nutzens der Neuroökonomie für die Ökonomie.

Die Befürworter der Neuroökonomie waren sich weitgehend einig, dass der globale Nutzen für die Ökonomie in der Öffnung der Black Box liegen wird, die das menschliche Gehirn in der ökonomischen Theorie darstellt (siehe beispielsweise Camerer et al. 2005b, 2004, Bernheim 2009, Reimann et al. 2011). Neuroökonomie würde diese Black Box öffnen, „much as organizational economics opened up the theory of the firm.“ (Camerer et al. 2004, 555) Jene Öffnung der Black Box hatte den Ökonomen nach Ansicht der Autoren neue Einblicke und Möglichkeiten beschert, die Neuroökonomie könnte das auch tun. Dieses Versprechen ist an sich schon recht hoch gegriffen, schon allein deshalb, weil gar nicht klar ist, ob sich diese Black Box überhaupt öffnen lässt und falls ja, ob das einen Nutzen für die Ökonomie bringt (Gul et al. 2008). Für manche trägt auch die Analogie mit der Theorie der Firma nicht (Bernheim 2009). Zudem ist das Versprechen der Öffnung der Black Box so oberflächlich formuliert, dass nicht von selbst evident wird, wie ein solcher Prozess aussehen könnte. Daher mussten die Befürworter der Neuroökonomie weiter ins Detail gehen, was den möglichen Nutzen der Neuroökonomie für die Ökonomie betrifft. Camerer, Loewenstein und Prelec zeigen in mehreren Veröffentlichungen verschiedene mögliche Wege auf, von denen wir zwei betrachten (Camerer et al. 2004, 573):

„[...] basic neuroeconomic research will ideally be able to link hypotheses about specific brain mechanisms (location, and activation) with unobservable intermediate variables (utilities, beliefs, planning ahead), and with observable behavior (such as choices). One class of fruitful tasks is those where some theories assume choice A and choice B are made by a common mechanism, but a closer neural look might suggest otherwise.“

Aus diesen und auch aus weiteren Ausführungen der Autoren wird leider überhaupt nicht klar, wieso das von Nutzen für die Ökonomie sein sollte oder was die

Verhaltensökonomien bei ihrer Arbeit mit diesen möglicherweise demnächst verfügbaren Erkenntnissen anfangen sollten. Ganz ähnlich verhält es sich mit dem darauffolgenden Beispiel (Camerer et al. 2004, 574):

„A third payoff from neuroscience is to suggest that economic choices which are considered different in theory are using similar brain circuitry. For example, studies cited above found that insula cortex is active when players in ultimatum games receive low offers, when people choose ambiguous gambles or money, when people see faces of others who have cooperated with them, and in players who are poor strategic thinkers. This suggests a possible link between these types of games and choices which would never have been suggested by current theory.“

Auch hier bleibt wieder die Frage offen, inwiefern diese Erkenntnisse nützlich für die Ökonomie sind, warum Ökonomen sie benötigen. Es wird aber klar, wie manche Ökonomen zu der Ansicht gelangten, die Neuroökonomie könnte die Probleme der Ökonomie nicht lösen.

Diese und andere Nutzenbeispiele haben gemeinsam, dass sie nicht hinreichend überzeugend und manchmal sogar überzogen waren. Kein Wunder also, dass die Verunsicherung auf allen Seiten groß war, wie nun mit der Neuroökonomie zu verfahren sei. Doch wie so oft kommt auch hier der Rat mit der Zeit. Seit den Hochzeiten der Debatte sind, wie gesagt, ungefähr zehn Jahre vergangen und einige neuroökonomische Studien und Arbeiten wurden durchgeführt und veröffentlicht.

Was also ist nun vom Glauben an den Nutzen der Neuroökonomie geblieben? Wie wir etwas detaillierter in Kapitel 6 sehen werden, scheint ein gewisser Realismus eingezogen zu sein, vor allem bei den eher neutral eingestellten Forschern beider Seiten. Die überzogenen Versprechen von damals wurden als solche identifiziert, teilweise von den Autoren selbst. Eine solche Vorgehensweise findet man heute kaum mehr, die Neuroökonomien scheinen daraus gelernt zu haben. Auch mit etwas Erfahrung im Gepäck wird der Nutzen der Neuroökonomie für ihre Elternwissenschaften heute pragmatischer eingeschätzt. Den Nutzen der Neuroökonomie, wie er sich heute darstellt, können wir in den beiden vorangegangenen Unterkapiteln 5.1 und 5.2 sehen. Dort haben wir besprochen, welchen Beitrag die Ökonomie beziehungsweise die Neurowissenschaften zur Neuroökonomie liefern und haben daraus einen Nutzen für die jeweils andere Wissenschaft abgeleitet:

Die Neurowissenschaftler können aus der neuroökonomischen Zusammenarbeit vor allem den Nutzen ökonomischer Modelle ziehen, sowie das Know-how, sie richtig anzuwenden. In den klassischen kognitiven Neurowissenschaften wird oft auf experimentelle Paradigmen aus der psychologischen Forschung des

20. Jahrhunderts zurückgegriffen. Da die Entwicklung leistungsfähiger Experimentparadigmen keine leichte Aufgabe ist, hat dieses Vorgehen grundsätzlich seinen Sinn. So können die Neurowissenschaftler von den Erfahrungen ihrer psychologischen Vorgänger profitieren. Andererseits holen sich die kognitiven Neurowissenschaftler damit nicht nur die Vor- sondern auch die Nachteile vieler psychologischer Modelle ins Labor. Nicht selten beruhen sie auf Konstruktbasis und sind dadurch nicht immer leicht operationalisierbar, wie in Abschnitt 5.1 am Beispiel des *Balloon Analogue Risk Task* zu sehen ist. Ökonomen und Neuroökonomien sehen die Modelle der Verhaltensökonomie zur Erforschung menschlicher Entscheidungsfindung klar im Vorteil. Sie sind in Jahrzehnten gut erprobt und decken ein großes Spektrum an verschiedenen Anwendungsmöglichkeiten ab, da sie grundsätzlich auf jede zu erforschende Situation anwendbar sind. Dazu noch mit klar strukturierten Situationen und Abläufen, die gewisse Wahrscheinlichkeiten beinhalten und sich klar operationalisieren lassen. Diesen Nutzen könnten kognitive Neurowissenschaftler aus der Neuroökonomie erhalten. Umsichtig angewandt könnte er die Entwicklung der kognitiv neurowissenschaftlichen Forschung zur Entscheidungstheorie fördern und beschleunigen, denn nicht nur könnten die Forscher die Vorteile der ökonomischen gegenüber psychologischen Modelle nutzen, es findet sich auch nicht für jede Forschungsfrage ein passendes experimentelles Paradigma aus der Psychologie, das sich für Studien mit fMRT, PET oder EEG übersetzen ließe, sodass ein eigenes erdacht werden muss, was, wie gesagt, eine nicht ganz triviale Aufgabe ist. Ein erfolgreiches Anwendungsbeispiel eines ökonomischen Experimentparadigmas in einer neurowissenschaftlichen fMRT-Studie haben wir in Abschnitt 5.2 gesehen.

Der mögliche Nutzen der Neuroökonomie für die Ökonomie wird hart diskutiert. Während die einen Ökonomen glauben, dass Ökonomie und Neurowissenschaften keine gemeinsame Basis haben und schon allein so unterschiedlichen Fragestellungen und Problemen nachgehen, dass die Neurowissenschaften keinen Nutzen für die Ökonomie haben oder haben werden, sind andere der Überzeugung, die Neuroökonomie könne die Präzision von Funktionen und Parametern in ökonomischen Standardmodellen erhöhen (Camerer et al. 2004, 574). Während noch über die wissenschaftstheoretische Vereinbarkeit von Ökonomie und Neurowissenschaften und über Brückengesetze diskutiert wird (siehe Kapitel 6), könnte der tatsächliche, verwertbare Nutzen der Neuroökonomie für die Ökonomie viel banaler liegen: In Abschnitt 5.2 betrachteten wir das Experiment von Zaki und Mitchell zu neuronalen Implementierungen gerechter Entscheidungen. In ihrer Studie gehen sie der Frage nach, warum Menschen in sozialen Entscheidungssituationen nicht selten fair oder gerecht handeln, genauer gesagt, warum Diktatoren im Diktatorspiel ihren Empfängern überhaupt Geld zuteilen.

Dass die Diktatoren im Spiel so handeln, lässt sich auch in behavioralen Untersuchungen feststellen. Gerade aber für die Frage nach dem Warum könnten bildgebende Studien aus den kognitiven Neurowissenschaften neue Daten und Perspektiven zur Diskussion beitragen, die die Verhaltensökonomien mit behavioralen Studien allein so leicht nicht reproduzieren könnten. Zaki und Mitchell beispielsweise bringen starke Evidenz dafür, dass das generöse Verhalten von Diktatoren nicht mit einem „warm glow“-Gefühl, wie sie es nennen (Zaki et al. 2011, 19762), zusammenhängt, das bei altruistischen Handlungen entsteht. Die Aktivierung mit Belohnung zusammenhängender Gehirnregionen beim Treffen gerechter Entscheidungen legt nahe, dass prosoziale Entscheidungen mit der Erfahrung subjektiver Belohnung einhergeht. Gerechte Entscheidungen hätten also ihren eigenen intrinsischen Wert.⁶ Andersherum schließen die Autoren aus ihren Aktivierungsdaten, dass nicht nur diejenigen aversiv auf ungerechte Entscheidungen reagieren, die von ihnen betroffen sind, sondern auch diejenigen, die sie verursacht haben (Zaki et al. 2011, 19763).

Diese Erkenntnisse werden die Verhaltensökonomie nicht revolutionieren, vermutlich ebenso wenig wie andere oder zukünftige Studien aus den kognitiven Neurowissenschaften. Aber sie können Hinweise auf die Handlungsmotivationen der Versuchspersonen liefern, die Ökonomen in ihren Laboren nicht so einfach erschließen können. Natürlich könnte man mit Gul und Pesendorfer und vielen anderen nun sagen, dass solche Erkenntnisse für die Ökonomie irrelevant sind und die ökonomischen Probleme nicht lösen. Andererseits könnten sie Verhaltensökonomien zu neuen Forschungsfragen und Studien inspirieren. Denn wer mehr weiß, dem fällt auch mehr ein.

In diesem Kapitel sind bereits einige Kritiken an der Neuroökonomie angeklungen. Auf dieser Basis werden wir uns nun in Kapitel 6 mit diesen und weiteren Kritiken beschäftigen und ihre Wirkung auf die Neuroökonomie betrachten.

⁶ Es ist zu beachten, dass Zaki und Mitchell prosoziales Verhalten als solches Verhalten definieren, das in ihrem Experiment bei einer Entscheidung derjenigen Person das Geld zuspricht, die den höheren Betrag erhält, gleichgültig, ob das nun die Versuchsperson im Scanner selbst ist, oder die andere, von der ein Foto eingeblendet ist. Das Merkmal, auf das es Zaki und Mitchell ankommt, ist, dass durch die Entscheidung der Versuchsperson der größtmögliche Gewinn realisiert wird, ob nun für die Versuchsperson selbst oder für die andere. Das heißt, es kann auch ein Gefühl der Belohnung auftreten, wenn die Versuchsperson sich selbst das Geld zugesprochen hat, sofern für sie auch der höhere Betrag vorgesehen war. Das unterscheidet sich insofern von einem *Warm Glow*, dass dieser nur bei altruistischem Verhalten auftritt, also in unserem Beispiel genau dann, wenn die Versuchsperson der anderen Person Geld zuspricht, egal, wer den höheren Betrag erhalten hätte.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.





Kritiken an der Neuroökonomie

6

In Kapitel 5 haben wir die neuroökonomische Zusammenarbeit und ihre möglichen Nutzen für ihre Mutterwissenschaften Verhaltensökonomie und kognitive Neurowissenschaften betrachtet. Diese Zusammenarbeit findet in der wissenschaftlichen Gemeinschaft aber nicht nur Anhänger, auch Kritiker tragen zu der nicht geringen Zahl an Publikationen zum Thema Neuroökonomie bei. Die Gründe, die sie vorbringen, sind vielzählig: Manche Wissenschaftler sehen grundsätzliche Probleme, die immer verhindern werden, dass kognitive Neurowissenschaften und Verhaltensökonomie je zu einer fruchtbaren Kooperation zusammenkommen; andere sehen Probleme, die sich eventuell mit der Zeit lösen ließen, im Augenblick aber eine Zusammenarbeit erschweren.

In diesem Kapitel werden wir die Kritiken, die gegen die Neuroökonomie vorgebracht werden, systematisieren und genauer untersuchen.¹ Welcher Art sind diese Kritiken, woher rühren sie, und könnten die angesprochenen Probleme gelöst werden? An ihren Kritiken muss sich die Neuroökonomie messen lassen und die Frage nach der Lösbarkeit kann die Zukunft dieser noch jungen Wissenschaft stark beeinflussen.

Um sich besser mit den Kritiken auseinandersetzen zu können, teile ich sie in zwei Kategorien: Die erste Kategorie, die wir in Abschnitt 6.1 betrachten werden, beinhaltet solche Kritiken, die nicht systematische Grundlagenproblematiken ansprechen, sondern vor allem die angewandte Zusammenarbeit von kognitiven Neurowissenschaftlern und Verhaltensökonomern betreffen. Hier geht es um überzogene Versprechungen, die Neuroökonomern machen oder in der Vergangenheit

¹ Die Reihe an Kritiken, die in diesem Kapitel vorgestellt werden, beansprucht keinesfalls Vollständigkeit, sicherlich gibt es weitere Kritiken. Um sie alle aufzunehmen, fehlt an dieser Stelle jedoch der Platz. Die vorgestellten Kritiken geben aber meines Erachtens einen guten Überblick über die Bandbreite der vorgebrachten Probleme.

gemacht haben, und die den Ruf der Neuroökonomie ungünstig beeinflusst haben. Zudem wird unzureichende technische Ausrüstung als Grund dafür genannt, dass neuroökonomische Erkenntnisse nicht belastbar sind. Auch neurowissenschaftliche Forschungsstandards geraten in die Kritik, da nach Ansicht mancher Ökonomen die Zahl der Probanden in neurowissenschaftlichen Experimenten zu niedrig sind, um repräsentative Schlussfolgerungen zuzulassen. Dieses Problem ist verbunden mit der Kritik, dass Neurowissenschaftler ihre Forschungsdaten nicht offenlegen, um Überprüfungen zuzulassen. In der fünften Kritik, die wir in dieser Kategorie betrachten werden, geht es um den zu wenig reflektierten Umgang der meisten Neurowissenschaftler mit *reverse Inference*. Wir werden sehen, wie diese Kritiken ausformuliert werden und wie ihnen von Seiten der Neuroökonomien begegnet wird, oder auch begegnet werden sollte. Bei dieser Art von Kritiken gibt es Hoffnung auf eine Lösung der Probleme.

Die zweite Kategorie, die in Abschnitt 6.2 behandelt wird, beinhaltet Kritiken grundlegender Art, die nicht zuerst die tägliche Arbeit der Wissenschaftler betreffen, sondern in den Eigenschaften der Mutterwissenschaften fußen. Hier sind mögliche Lösungen nicht so leicht gefunden. Es geht dabei um Kritiken, die vor allem von den Ökonomen Faruk Gul und Wolfgang Pesendorfer geäußert wurden. Sie beklagen, dass Neuroökonomien ökonomische Modelle und Theorien missverstehen und sie deshalb als schlecht geeignet ansehen. Ihrer Meinung nach sind ökonomische Definitionen und Modelle aber erfolgreich und wertvoll, weil sie sind, wie sie sind. Änderungen, die Neuroökonomien vorschlagen, sehen sie als nicht hilfreich. In einer zweiten Kritik argumentieren sie für die Irrelevanz neurowissenschaftlicher Erkenntnisse für die Ökonomie, was sie auf deren unterschiedliche Fragestellungen, Ziele, Instrumente und Datenquellen zurückführen. Da die Ökonomie keine Angaben über Gehirnfunktionen macht, können nach Gul und Pesendorfer ökonomische Modelle nicht durch Gehirndaten widerlegt werden. Diese Haltung wird unter anderem von Ökonomen als unnötig isolierend angesehen. Wir werden sehen, dass Guls und Pesendorfers Kritiken an der Neuroökonomie auf ihrer behavioristischen Konzeption von der Ökonomie als Wissenschaft beruht und unvereinbar ist mit der vollständig verschiedenen neuroökonomischen Konzeption der Ökonomie. Diese Differenzen haben fundamentale und unüberwindbare Konsequenzen für die Zusammenarbeit von Ökonomen und Neurowissenschaftlern, die sich nicht mit besserer Kommunikation und gemeinsamen Forschungsstandards lösen lassen.

Guls und Pesendorfers Behaviorismus wird allerdings bei weitem nicht von jedem Ökonomen geteilt und steht selbst seit Jahrzehnten in der Kritik. Doch auch wenn man mit ihrer Einstellung nicht übereinstimmen mag, so bleiben doch

Kritiken an der Neuroökonomie, die auch Verhaltensökonominnen mit Kooperationsinteresse nicht einfach aus dem Weg räumen können. Diese Probleme werden in der Wissenschaftstheorie seit Jahrzehnten immer wieder diskutiert, wir werden sie in Abschnitt 6.3 näher betrachten. Hier geht es um die Probleme, die bei einer Kooperation so verschiedener Wissenschaften wie Ökonomie und Neurowissenschaften aufkommen, also um Reduktionismus mit all seinen Schwierigkeiten und um Inkommensurabilität, wie Thomas Kuhn sie darlegt. Gerade die Frage, ob neurowissenschaftliche Erkenntnisse dabei hilfreich sind, Modelle zu entwickeln, die nicht einzelne Neuronen oder Neuronenverbände zum Inhalt haben, sondern das Verhalten des Menschen, wurde im 20. Jahrhundert bereits heftig diskutiert, unter anderem von Jerry Fodor, dessen Arbeit wir besonders betrachten werden. Das sind für Neuroökonominnen heute relevante Themen, die sie kennen sollten, um sowohl mit Kooperation als auch mit Kritik erfolgreich umgehen zu können.

6.1 Die prozeduralen Kritiken

Im Folgenden werden wir uns drei Kritiken ansehen, die sich mit der täglichen wissenschaftlichen Arbeit, den verwendeten Instrumenten und der Einstellung der Neuroökonominnen zur Leistungsfähigkeit der Neuroökonomie befassen. Beginnen wir mit Letzterem.

6.1.1 Die hohen Versprechungen der Neuroökonominnen

Diese Kritik, die immer wieder und vor allem von Ökonominen geäußert wird, beanstandet die großen Versprechungen über die potenzielle und zukünftige Leistungsfähigkeit der Neuroökonomie.

Vor allem zu Beginn der neuroökonomischen Kooperation warben ihre Anhängerinnen damit, die großen Fragen und Probleme der Ökonomie und Psychologie künftig lösen zu können und eine neue, umfassende Theorie über den Prozess der menschlichen Entscheidungsfindung aufzustellen. Besonders Colin Camerer, George Loewenstein und Drazen Prelec, die zusammen mehrere Publikationen zum Thema verfassten, zeigten sich so begeistert von der Neuroökonomie, dass sie deren Potenzial sehr optimistisch einschätzten. Sehen wir uns ein Beispiel an:

„In a strict sense, all economic activity must involve the human brain. Yet, economics has achieved much success with a program that sidestepped the biological and cognitive sciences that focus on the brain, in favor of the maximization style of classical

physics, with agents choosing consumption bundles having the highest utility subject to a budget constraint, and allocations determined by equilibrium constraints. [...] Of course these economic tools have proved useful. But it is important to remember that before the emergence of revealed preference, many economists had doubts about the rationality of choice. [...] Jevons (1871) wrote, 'I hesitate to say that men will ever have the means of measuring directly the feelings of the human heart. It is from the quantitative effects of the feelings that we must estimate their comparative amounts.' The practice of assuming that unobserved utilities are revealed by observed choices – revealed preference – arose as a last resort, from skepticism about the ability to 'measure directly' feelings and thoughts. But Jevons was wrong. Feelings and thoughts *can* be measured directly now, because of recent breakthroughs in neuroscience. If neural mechanisms do not always produce rational choice and judgment, the brain evidence has the *potential* to suggest better theory." (Camerer et al. 2004, 555–556)

Camerer, Loewenstein und Prelec sehen den Grund dafür, dass die Ökonomie *revealed Preference* als grundlegendes Modell verwendet, in der letzten Rettung der Ökonomen vor dem Problem, dass sich Gefühle und Gedanken ökonomischer Agenten nicht direkt messen ließen. Das Gehirn wird in der Folge als Black Box behandelt, wodurch es nicht nötig ist, zu messen, was darin vorgeht. Das Modell von *revealed Preference* ist also laut Camerer, Loewenstein und Prelec lediglich eine Art Behelfsmittel, da die eigentlich erstrebte Einsicht in Gefühle und Gedanken der ökonomischen Agenten nicht zur Verfügung stand. Dieser wissenschaftshistorisch gewagten Hypothese würden einige moderne Ökonomen sicher nicht zustimmen (vgl. z. B. Gul et al. 2008). Für uns an dieser Stelle vor allem interessant ist aber die Behauptung der Autoren, Neurowissenschaftler seien nun aufgrund kürzlicher Durchbrüche in ihrer Wissenschaft dazu in der Lage, Gefühle und Gedanken direkt zu messen. Somit könnte der alte Missstand behoben und endlich ein Modell in der Ökonomie eingeführt werden, das die Befunde aus der Hirnforschung berücksichtigt.

Nun zeigen sich viele Wissenschaftler zwar gerne selbstbewusst, wenn es um die Leistungsfähigkeit ihrer Wissenschaft geht, aber Gefühle und Gedanken von Menschen direkt messen zu können, ist doch überzogen. Die Ökonomen Camerer, Loewenstein und Prelec sind sich aber im Jahr 2004 nicht nur sicher, dass das bereits möglich ist, sie sind auch überzeugt, dass diese Messungen das Potenzial haben, den Ökonomen eine bessere Theorie zu unterbreiten als diejenige, die seit vielen Jahrzehnten als einer der Grundpfeiler der Ökonomie gilt. Dass solche Ansichten vor allem unter Ökonomen Kritik hervorrufen würden, war absehbar. Allen voran Gul und Pesendorfer kritisieren solche Aussagen als nicht zutreffend und für die Ökonomie wenig hilfreich (Gul et al. 2008, 33).

Vielleicht haben auch Camerer und seine Koautoren eingesehen, dass ihre Aussage überzogen war, denn ein Jahr nach dem oben zitierten Artikel veröffentlichten sie einen weiteren, ganz ähnlichen, in dem sie Jevons' Zitat noch einmal verwendeten, ihre darauffolgenden Ausführungen aber zurückhaltender formulierten (Camerer et al. 2005, 10):

„But now neuroscience has proved Jevons's pessimistic prediction wrong; the study of the brain and nervous system *is* beginning to allow direct measurement of thoughts and feelings. These measurements are, in turn, challenging our understanding of the relation between mind and action, leading to new theoretical constructs and calling old ones into question.“

In dieser Version sind die Neurowissenschaften nach Ansicht der Autoren nicht bereits in der Lage, Gefühle und Gedanken direkt zu messen. Vielmehr fängt diese Entwicklung gerade erst an. Dass die direkte Messung von Gefühlen und Gedanken mithilfe der Hirnforschung künftig möglich sein wird, daran halten sie jedoch fest. Außerdem hat sich die Darstellung der neuroökonomischen Einflussmöglichkeiten auf ökonomische Grundlagen geändert. Hier sprechen die Autoren nicht mehr von einem direkten Weg der Erkenntnisse aus den Neurowissenschaften in die Ökonomie. Hier gehen sie einen Umweg über eine Veränderung unseres Verständnisses des Zusammenhangs von Geist und Handlung durch neue Erkenntnisse aus den Neurowissenschaften. Dieses neue Verständnis wiederum wird den Autoren zufolge zu neuen theoretischen Konstrukten führen, die die alten (ökonomischen) in Frage stellen.

Zwar klingt das immer noch so, als wollten Camerer und seine Kollegen mithilfe der Neuroökonomie eine Kuhn'sche Revolution in der Ökonomie anzetteln, ist aber bereits ein Beispiel dafür, dass die hohen Versprechungen nach den Anfangsjahren der Neuroökonomie flacher werden. Die Zeit der großen Motivationspublikationen, die, mitunter durch spektakuläre Versprechungen, möglichst viel Aufmerksamkeit generieren sollten, ist inzwischen vorbei (die genannten Artikel von Camerer und seinen Koautoren stammen immerhin bereits aus den Jahren 2004 und 2005). Das sieht offenbar auch Camerer selbst so, denn in einem Sammelbandbeitrag aus dem Jahr 2008 schreibt er (Camerer 2008a, 44):

„This chapter develops my latest view about grounding economic choice in neural details. This perspective is developing rapidly. As a result, viewpoints expressed only a couple of years ago are updated and informed by the latest data and perspective on methods. Revision of viewpoints, particularly the details of language and its sweep, is common and desirable in empirical science as new data and methods arrive. Early neuroeconomics papers that describe ideas and potential discoveries [e.g., Camerer,

Loewenstein, and Prelec, 2005] should therefore not be viewed as logical conclusions derived from mathematical analysis. These early neuroeconomics papers should be read as if they are speculative grant proposals that conjecture what might be learned from studies that take advantage of technological advances.“

Es scheint, als hätten die Neuroökonominnen die Kritiken zum Thema überzogene Versprechen verstanden und erkannt, dass sie diese Praxis nicht weiterverfolgen sollten. Was die Neuroökonominnen jetzt liefern wollen, sind neue Erkenntnisse, die in täglicher Forschungsarbeit gewonnen werden und die Neuroökonomie nach vorne bringen sollen.

Ein weiteres Beispiel, das auch in diese Kategorie gehört, ist eine ambitionierte Formulierung der Ziele der Neuroökonomie. In einem Beitrag zu einem neuroökonomischen Sammelband aus dem Jahr 2009 schrieb der Neurowissenschaftler Paul Glimcher (Glimcher 2009c, 503): „The goal of neuroeconomics is an algorithmic description of the human mechanism for choice.“ Leider erläutert er nicht, was er unter einer „algorithmic description“ genau versteht, was gegen den Vorwurf überzogener oder schlecht formulierter Versprechungen nicht hilfreich ist. So bleibt den Lesern die Interpretation selbst überlassen. Offenbar sieht Glimcher den menschlichen Entscheidungsprozess als Verfahren, das sich in einzelnen Schritten beschreiben und durchführen lässt. Diese Ansicht kann man durchaus teilen und vielleicht haben die kognitiven Neurowissenschaften auch in mehr oder weniger naher Zukunft genügend Erkenntnisse über die Arbeit des menschlichen Gehirns gewonnen, um diese einzelnen Schritte identifizieren zu können. Dann bleibt aber noch die Frage, ob es *den* menschlichen Entscheidungsprozess überhaupt gibt. Das klingt, als basierten alle möglichen Entscheidungen auf einem immer gleich ablaufenden Prozess. Dabei werden Hirnprozesse doch gerade von den Kognitionswissenschaften in affektive und kognitive unterschieden, wodurch sich unterschiedliche Wege zu Entscheidungen ergeben, von unterschiedlichen Wegen von Informationen in und aus dem Gehirn ganz abgesehen. Wie sollen all diese feinen Unterschiede algorithmisch beschrieben werden, ohne eine ganze Bibliothek an Beschreibungen zu produzieren? Diese Idee ist jedenfalls erklärungsbedürftig und Glimcher hätte dies weiter ausformulieren sollen. So bleibt es eine schlecht beschriebene und missverständliche Zielformulierung, von der die Leser nicht wissen, wie sie einzuordnen ist. Wollen sich die Neuroökonominnen des Vorwurfs erwehren, sie hätten die Ökonomie falsch verstanden und schätzten die Leistungsfähigkeit der kognitiven Neurowissenschaften weit zu optimistisch ein (Gul et al. 2008; Harrison 2008a), sind solche Publikationen nicht hilfreich. Zielformulierungen und Versprechungen wie

diese sind nicht nur selbst schnell kritisiert, sie lassen auch Zweifel daran aufkommen, ob sich die Neuroökonomie nicht übernimmt und von vornherein zum Scheitern verurteilt ist. Dieses Bild sollten die NeuroökonomInnen vermeiden, da ihnen sonst die Unterstützung der Wissenschaftler wegbrechen könnte.

Glimchers Formulierung ist sicherlich ein einfacher Angriffspunkt. Allerdings habe ich diese Zielsetzung in keiner weiteren Publikation, früher oder später, wiedergefunden. Entweder hat er selbst gesehen, dass sie einen vermeidbaren Angriffspunkt bietet, oder er hat sie selbst gar nicht so ernst genommen. Das würde auch erklären, warum er sie (auch später) nicht weiter erläutert hat. Das ist zwar kein Beispiel für gute wissenschaftliche Publikationspraxis, untermauert aber den Verdacht, dass es sich bei Glimchers Aussage um seine persönliche Zielformulierung handelt und nicht um eine in Zement gegossene Zielsetzung einer neuroökonomischen Gemeinschaft.

Doch selbst wenn mehrere NeuroökonomInnen Glimchers Traum von einer algorithmischen Beschreibung der menschlichen Entscheidungsmechanismen teilen, wäre das ein so starker Kritikpunkt, dass die neuroökonomische Zusammenarbeit aufgegeben werden müsste? Nicht unbedingt. Wenn die Neuroökonomie ein schlecht formuliertes Großziel nicht erreicht, folgt daraus allein nicht, dass sie nicht durch das Erreichen von Zwischenzielen neue, für die Mutterwissenschaften relevante Erkenntnisse liefern kann. Das große Ziel eines Projekts erhält für gewöhnlich viel Aufmerksamkeit, aber auch die Fortschritte, die auf dem Weg dorthin gemacht werden, können selbst wertvoll sein. Kritiker sollten daher vorsichtig damit sein, die gesamte Neuroökonomie zu verwerfen, weil frühe Versprechungen von Zielen oder revolutionären Beiträgen zur Ökonomie überzogen oder schlecht formuliert waren. Ökonom Glenn Harrison spricht von „marketing bloopers“ (Harrison 2008a, 304), die mit einem akademischen Hype daherkommen und die man nicht zu ernst nehmen sollte und schließt (Harrison 2008a, 339):

„We can put the academic marketing of neuroeconomics aside. It almost seems unfair to put some of those claims on display, but [...] they have taken hold in many quarters as knowledge claims when they are just ‘chloroform in print’.“

Kurzum, die NeuroökonomInnen haben in diesem Fall ihre Kritiker zur Kenntnis genommen und halten ihre anfänglichen Versprechen inzwischen selbst für überzogen. Daher sollte dieser Kritikpunkt kein mehr Grund sein, neuroökonomische Ideen rundheraus abzulehnen.

6.1.2 Unzureichende Technik

Die zweite Kritik aus dieser Gruppe beschäftigt sich mit den technischen Instrumenten, die kognitive Neurowissenschaftler für ihre Experimente verwenden. Es wird kritisiert, dass Neuroökonomien die Messtechniken als leistungsfähiger ansehen, als sie tatsächlich sind. Das bedeutet erstens, dass die in Experimenten gewonnenen Daten nicht die Rückschlüsse zulassen, die viele Forscher aus ihnen ziehen möchten und zweitens folgt für die Kritiker daraus, dass die veröffentlichten Experimentdaten und -ergebnisse nicht verlässlich genug sind.

Vor allem die Positronenemissionstomographie (PET) und die funktionelle Magnetresonanztomographie (fMRT) stehen in der Kritik. Wie wir weiter oben (Kap. 5) bereits betrachtet haben, basieren sie auf der Ermittlung hämodynamischer Veränderungen, also Veränderungen im Blutfluss zu aktivierten Neuronen. Da diese Veränderungen eine gewisse Zeit benötigen, um aufzutreten und wieder abzuklingen, haben die hämodynamischen Techniken den Nachteil einer relativ groben zeitlichen Auflösung von mehreren Sekunden. Das bedeutet, dass die Aktivierung, die durch einen erhöhten Blutfluss angezeigt wird, tatsächlich bereits vor einigen Sekunden geschehen ist. Das macht es schwieriger, die Aktivierungen den kognitiven Prozessen zuzuordnen, die in der Experimentaufgabe gefragt sind. Zuschreibungen von Aktivitätsmustern zu bestimmten kognitiven Leistungen könnten also zu ungenau oder gar fehlerhaft ausfallen.

Generell gelten die hämodynamischen Messtechniken aber als relativ genau, was die räumliche Auflösung betrifft, die sich im Rahmen weniger Millimeter bewegt. Doch auch hier sehen Kritiker die Gefahr falscher Zuordnungen (Cabeza et al. 2000, 2) durch die nach den Experimentdurchläufen notwendige Datenvorverarbeitung. Vor allem die Glättung kann dazu führen, dass die räumliche Auflösung zu grob wird, um Organisationen auf neuronalem Level beobachten zu können.²

Darüber hinaus stellt sich für manchen Kritiker die Frage, ob die in Experimenten gewonnenen Aktivierungsmuster grundsätzlich aussagekräftig sind. Wie bereits oben beschrieben³, liegen Messwerte für aktive Neuronen nur sehr wenig über denen für gerade nicht oder wenig aktive Neuronen, sodass leicht falsch positive Aktivierungen gemeldet werden können. Auch die Filterung von sogenanntem *Noise* ist nicht einfach zu bewerkstelligen, wie wir gesehen haben. Daten aus einem fMRT könnten also tatsächlich nicht genau genug sein, um Schlüsse über Aktivierungsmuster zuzulassen, meinen Kritiker (Ortmann 2008; Thirion

² Für nähere Erläuterungen zur Glättung von Experimentdaten siehe oben, Abschn. 5.2.

³ Siehe oben, Abschn. 4.3.1.

et al. 2007). Der Ökonom Andreas Ortmann (Ortmann 2008, 441) bringt seine Ansichten auf den Punkt: „The colourful pictures suggest more precision than there is and hence invite overselling and hyperbole.“

Einen anderen Aspekt des Problems mit den Messtechniken zeigen die Neurowissenschaftler Roberto Cabeza und Lars Nyberg auf: Sie geben zu bedenken, dass hämodynamische Messtechniken nicht nur technische sondern auch methodische Schwächen besitzen, die die Forschung nicht ignorieren sollte (Cabeza et al. 2000, 3):

„[...] all these techniques have a general limitation: They identify brain regions associated with a certain cognitive process, but do not provide information regarding the functional relations between these regions.“

Daten aus einem fMRT-Computer zeigen die einzelnen Aktivierungen verschiedener Voxel an. Ob diese Aktivierungen aber in irgendeinem Zusammenhang stehen, und wenn ja, in welchem, lässt sich nur anhand dieser Daten nicht klären. Dabei sind Einsichten in die Zusammenhänge zwischen den Aktivierungen verschiedener Hirnregionen essentiell für ein Verständnis der Abläufe kognitiver Prozesse.

Zudem geben Cabeza und Nyberg (Cabeza et al. 2000, 2) zu bedenken, wie wichtig das Experiment-Design bei bildgebenden Verfahren ist. Von ihm hängt ab, ob die gesammelten Daten aussagekräftig sind oder nicht. Die dabei in frühen Zeiten oft verwendete Subtraktionsmethode (s. Abschn. 5.2) halten sie für nicht geeignet, aufgabeninduzierte Aktivierungen korrekt herauszufiltern.

Helfen uns die vielen fMRT-Studien also nicht, kognitiven Prozessen auf die Spur zu kommen? Sind die Aktivierungen, die so anschaulich in bunten Farben dargestellt werden, tatsächlich nur verschwommene und unzuverlässige Einblicke in die Durchblutung des menschlichen Gehirns und daher nicht für wissenschaftlichen Erkenntnisgewinn geeignet? Ganz so schlimm ist es sicher nicht. Sicher hat jede Technik ihre Nachteile und Schwächen. Diese sind den Neurowissenschaftlern und Psychologen auch seit Jahrzehnten bewusst, die meisten dieser Art von Kritiken stammen immerhin aus den Anfangsjahren der Verwendung bildgebender Verfahren. Heute werden bereits in der Einführungsliteratur in das experimentelle neurowissenschaftliche Arbeiten die technischen Nachteile und systematischen Probleme der einzelnen Messtechniken sowie Lösungswege aus der gängigen Praxis angesprochen.⁴

⁴ Siehe beispielsweise Hermey et al. 2010.

Das Problem mit der relativ schlechten zeitlichen Auflösung ist charakteristisch für Verfahren, die auf dem BOLD-Effekt beruhen und kann daher nicht gänzlich ausgeschaltet werden. Doch seit Einführung der Scanner in die tägliche medizinische und neurowissenschaftliche Arbeit in den 1990er Jahren hat sich die Technik der Scanner, Computer und Softwares stark verbessert. Arbeitete man zu Beginn noch mit 1,5 T-Scannern, sind heute 12 T-Scanner keine Seltenheit. Das hat die Genauigkeit der Messergebnisse in den vergangenen zwanzig Jahren verbessert. Ob eine weitere Erhöhung der Feldstärke zu noch genaueren Ergebnissen führen wird, darüber wird noch debattiert, aber generell ist eine weitere Verbesserung der Technik in Zukunft zu erwarten. Mit der Verbesserung der Technik hat sich auch das Problem etwas entschärft, dass die Werte aktivierter Neuronen nur geringfügig höher sind als die nicht aktivierter Neuronen. Feinere Messungen erkennen die Unterschiede besser.

Das Gleiche gilt für die Frage nach der räumlichen Auflösung. Auch sie wurde durch die Verbesserung der Technik erhöht. Zudem ist die Datenvorverarbeitung heute ein durchdachter, standardisierter Prozess, der von Software gesteuert wird, die von professionellen Unternehmen zu diesem Zweck geschrieben wurde. Vermutlich war das zu Beginn der fMRT-Studien weniger der Fall, doch werden heute beispielsweise nur noch zwei Standardkoordinatensysteme für die Normalisierung (also die Übertragung der Daten in eine genormte Gehirnschablone⁵) verwendet. Das schränkt die Variabilität der Normalgehirne zwischen Studien stark ein, was eines der kritisierten Probleme war. Auch für die statistische Analyse von Aktivierungsdaten zum intra- und intersubjektiven Vergleich gibt es inzwischen weltweit angewandte Standardverfahren. Dabei wird zwar kritisiert, dass das meistgenutzte, automatische Verfahren, die *voxel-based Morphometry*, nicht so akkurat ist wie das ältere, halbautomatisierte Verfahren, bei dem manuell alle Voxel einer interessierenden Region ausgewählt werden müssen, bei jedem Probanden und jedem Scannerdurchlauf. Die Menge an Daten, die moderne fMRT-Computer produzieren, würde jedoch eine Menge an manuell durchzuführenden Prozessen bedeuten, die die Anwendung des halbautomatischen Verfahrens impraktikabel macht. Die automatische Analyse bietet dagegen eine oft angewandte, daher erprobte Alternative, die in den vergangenen Jahrzehnten weiterentwickelt wurde und die sich sicherlich auch zukünftig weiter verbessern wird. Inzwischen wird sie auch seit einigen Jahren in der medizinischen Diagnostik eingesetzt, was nicht ohne Genehmigungsverfahren eingeführt wurde. So schlecht scheint es also um ihre Akkuratessse nicht bestellt zu sein.

⁵ Siehe Abschnitt 5.2 für weitere Erläuterungen.

Generell sind Kritiken, die sich auf die Technik der Messinstrumente bezieht, eher wenig spannend. Technik wird ständig verbessert, bereits jetzt sind die bildgebenden Verfahren leistungsfähiger als in ihren Anfangsjahren und werden auch in Zukunft leistungsfähiger werden. Kritik an der Technik ist heute schwächer als damals und morgen vielleicht schon gar nicht mehr relevant. Interessanter sind da die genannten Bedenken von Cabeza und Nyberg, die sich mit methodischen Problemen bei Studien mit bildgebenden Verfahren befassen. Bildgebende Verfahren können lediglich Bilder von Aktivierungszuständen liefern, aber keine Aussage über funktionale Zusammenhänge aufzeigen. Auch dieses Problem lässt nicht gänzlich ausräumen, aber abschwächen. In den vergangenen zwanzig Jahren haben neurowissenschaftliche und psychophysiologische Forscher viel Erfahrung im Umgang mit bildgebenden Verfahren gesammelt. Sommer (Sommer 2010, 240) berichtet, dass zu Beginn der bildgebenden Studien noch gar nicht viel darüber bekannt war, wie genau das gemessene Signal zustande kam oder welche neuronalen Prozesse ihm zugrunde liegen. Hinzu kam, dass die anfänglichen Studien nicht sehr theoriebasiert waren und trotzdem hohe Publikationszahlen bekamen. Es kamen daher bald Vorwürfe auf, die kognitiven Neurowissenschaftler produzierten bedeutungslose Farbbilder, die sie nicht einmal selbst interpretieren konnten; sogar Vergleiche mit der Phrenologie wurden laut. Diese Vorwürfe konnten in der wissenschaftlichen Gemeinschaft inzwischen entschärft werden. Nicht nur sind die Grundlagen des BOLD-Effekts und der Vielzahl an zellphysiologischen Veränderungen bekannt, die das Signal im fMRT ausmachen, auch sind die Studien heute in ihrem wissenschaftlichen Standard und Anspruch gestiegen.

Natürlich können fMRT-Bilder keine Abhängigkeiten oder Kausalitäten aufzeigen, aber das allein ist kein Problem. Problematisch ist es, wenn trotzdem Schlüsse aus fMRT-Studien allein gezogen werden, die Aussagen über funktionelle Zusammenhänge machen. Die angesprochenen Erfahrungen im Umgang mit bildgebenden Verfahren und die Standardisierung von Studien hat dazu geführt, dass sich die neurowissenschaftliche Gemeinschaft der Vor- und Nachteile dieser Verfahren bewusst ist und heute anders an Studien herangeht als noch vor zwanzig Jahren. Dabei hat sich eben auch herausgestellt, dass es für die Erforschung bestimmter Fragen sinnvoll sein kann, mehr als ein Messinstrument für eine Studie zu verwenden. Daten von bildgebenden Verfahren können beispielsweise erweitert werden um Daten aus Läsionsstudien. Das kostet Zeit und Geld, erhöht aber die Leistungsfähigkeit der Studien. Allerdings können auch mehr Experimente mit mehr Messinstrumenten diesen grundlegenden

Nachteil von fMRT-Geräten nicht beheben. Die Scanner können keine funktionellen Organisationen im Gehirn ermitteln und dürfen dafür auch nicht verwendet werden.

Die gewachsene Erfahrung im Umgang mit fMRT-Studien hilft auch, das Problem mit der Subtraktionsmethode zu beheben: Heute sind sich die Forscher sicher, dass sie für den Einsatz in fMRT-Studien nicht geeignet ist und haben andere Methoden entwickelt, den durch die Experimentaufgabe ausgelösten Aktivierungen auf die Spur zu kommen. Beispiele dafür sind das in Abschnitt 5.2 besprochene faktorielle oder parametrische Experimentdesign. Dieses Problem besteht heute also nicht mehr und mit weiterer Erfahrung werden auch hier noch weitere Ideen auf den Tisch kommen.

Insgesamt können die systematischen Nachteile und Grenzen der bildgebenden Verfahren ein Problem und Anlass zu Kritik sein, wenn sich die Forscher dieser Grenzen nicht bewusst sind oder sie bewusst überschreiten. Eine grundlegende Aufklärung über die Funktionsweise und methodischen Stärken und Schwächen der Instrumente ist ohne Frage geboten, am besten gleich während des Studiums. Gleichzeitig sind Kontrollen untereinander in der wissenschaftlichen Gemeinschaft wichtig, um ungestützte Schlussfolgerungen oder inadäquate Experimentdesigns nicht zu fördern. Das bringt uns zum nächsten Set an Kritiken, das sich mit den wissenschaftlichen Methoden und Standards beschäftigt, die Neurowissenschaftler in ihren Studien befolgen.

6.1.3 Neurowissenschaftliche Forschungsstandards

Neben den Bedenken, die die Technik betreffen, fragen vor allem Ökonomen nach der Qualität der Verfahren und Standards, die in neuroökonomischen Experimenten Anwendung finden.

Die Kritik betrifft die Anzahl an Probanden, die für gewöhnlich für Studien mit bildgebenden Verfahren, vor allem fMRT, rekrutiert werden. Harrison (Harrison 2008a, 311) hat mehrere Studien betrachtet und kommt zu dem Schluss: „It is common to see studies where the sample size is less than 10, and rarely does one see much more than a dozen or so.“ Es ist nicht erstaunlich, dass er als Verhaltensökonom das als eine geringe Zahl empfindet, arbeiten ökonomische Experimente doch meist mit einer viel größeren Zahl an Versuchspersonen. Die Frage, die sich daher vor allem Verhaltensökonominnen stellen, ist, ob sich aus einer so geringen Zahl an Versuchspersonen überhaupt Daten gewinnen lassen, die Schlussfolgerungen zulassen, wie Neuroökonominnen sie ziehen? Könnten nicht kleine Unterschiede in den funktionellen Aktivierungsmustern so übersehen

oder unterschätzt werden? (Vgl. Ortmann 2008, 438–439). Nicht selten reagieren Neuroökonominnen auf diese Kritik mit der Aussage, dass mehr Versuchspersonen keinen positiven Effekt auf die Schlussfolgerungen haben (vgl. Bhatt et al. 2005, 432). Tatsächlich findet sich eine ähnlich lautende Einschätzung auch in Lehrbüchern zum Experimentieren mit bildgebenden Verfahren (Sommer 2010, 254–255):

„Eine letzte Frage bei der Planung einer fMRT-Studie betrifft die notwendige Stichprobengröße. Aufgrund der relativ hohen Kosten einer fMRT-Messung möchte man diese natürlich so gering wie möglich halten, ohne Gefahr zu laufen, einen vorhandenen Effekt nicht nachweisen zu können. Bei einer Größe von 20–25 Personen erreichen sowohl die Reliabilität, also die Verlässlichkeit der Messung, als auch die Effektstärken ein befriedigendes Maß, die mit größeren Stichproben nur noch langsam zunehmen.“

Eine Zahl an Versuchspersonen, wie Ökonomen sie aus ihrer eigenen experimentellen Forschung gewohnt sind, scheint also in Studien mit bildgebenden Verfahren kein Segen zu sein. Beachtet man aber die im Zitat empfohlene Anzahl von 20 bis 25 Versuchspersonen, so übersteigt sie die von Harrison berichteten Zahlen von zehn oder selten mehr als ein Duzend deutlich. Die Studie von Zaki und Mitchell (Zaki et al. 2011), die wir in Kapitel 5 betrachtet hatten, startete mit 19 Teilnehmern, wobei die Daten von 4 davon nicht modelliert werden konnten, sodass Analyse und Diskussion auf den Daten von 15 Versuchspersonen beruhen. Andere Stichproben zeigen zum Teil zwischen 14 und 19 Probanden.⁶ Es finden sich aber auch Studien mit 20 oder 30 Versuchspersonen.⁷ Sind das nun ausreichend viele, damit Reliabilität und Effektstärken der Messungen ein befriedigendes Maß erreichen, wie Sommer es fordert? Es stellt sich die Frage, wie Sommer auf die Empfehlung von 20 bis 25 Versuchspersonen kam. Der zitierte Text gibt darüber keine weitere Auskunft. Auf Anfrage erläutert Sommer,⁸ dass frühere Studien tatsächlich teilweise nur sechs Teilnehmer umfassten, was aus seiner Sicht Stichprobenfehler sehr wahrscheinlich macht. Seine Empfehlung von 20 bis 25 Versuchspersonen beruht zum einen auf einzelnen Studien, die unter anderem durch Simulationen solche Werte nahelegten. Zum anderen beschreibt

⁶ Siehe beispielsweise Rao et al. 2008 mit ihrer in Kapitel 3 besprochenen BART-Studie oder Moll et al. 2006.

⁷ Siehe etwa Tricomi et al. 2010 oder Chang et al. 2011.

⁸ E-Mail-Verkehr mit Sommer, der zum Zeitpunkt des Entstehens dieses Textes am Universitätsklinikum Hamburg-Eppendorf tätig ist.

er die Empfehlung als eine Art gewachsene und in der wissenschaftlichen Fachgemeinschaft akzeptierte Größe. Dabei spielen die hohen Kosten für Messungen in Scannern eine Rolle. Auch wenn sie theoretisch kein Faktor bei der Festlegung der Probandenzahl sein sollten, sind sie es in seinen Augen wegen der begrenzten Mittel in der Praxis aber durchaus. Erfahrungsgemäß, sagt Sommer, sollten große, bedeutende Effekte bei Studien mit 25 Versuchspersonen aufgetaucht sein. Tatsächlich beobachtet er aktuell einen Trend zu größeren Stichprobengrößen, was er auch im Zusammenhang mit der Replikationskrise der jüngsten Zeit sieht. Dazu wurden gerade einige Studien veröffentlicht, die die Auswirkungen der Stichprobengröße auf die Effektstärke und Leistungsfähigkeit von fMRT-Studien untersuchen. Eine solche Studie von Geuter et al. (Geuter et al. 2018) untersucht das auf Grundlage der Daten von knapp 500 Probanden, die aus dem Human Connectome Project (HCP)⁹ stammen. Auch die Autoren berichten von großen Diskussionen in den vergangenen Jahren, inwiefern Studien mit kleinen Stichprobengrößen die Reliabilität neurowissenschaftlicher Studien unterminieren. Für ihre eigene Studie ziehen sie funktionelle Daten aus mehreren Tasks des HCP heran. Die ursprünglichen Studien des HCP hatten die Probanden Aufgaben zum Arbeitsgedächtnis, motorische Aufgaben, Emotionsverarbeitung und Glücksspiel in fMRT-Scannern durchführen lassen. Mit den Daten von mindestens 471 Probanden¹⁰ als Population untersuchten Geuter et al. die Fähigkeit kleiner Stichprobengrößen, die Effekte, die in der Gesamtpopulation beobachtet wurden, ebenfalls zu erfassen. Dazu nahmen sie $K = 100$ verschiedene Stichproben aus der Gesamtpopulation mit den Größen: $N = 10, 20, 40, 60, 80$ und 100 . Diese Größen sollten typische Stichprobengrößen von fMRT-Studien nachempfinden. Für die verschiedenen Experimentaufgaben (Arbeitsgedächtnis, Motorik, Emotionsverarbeitung und Glücksspiel) wurde über die festgelegten Stichprobengrößen berechnet, welche Anzahl an Probanden mindestens nötig ist, um eine empirische Leistungsfähigkeit von 50%, 80%, 90% und 95% zu erreichen. Die empirische Leistungsfähigkeit wurde veranschlagt als der Bruchteil von $K = 100$ Wiederholungen der Stichprobengröße N , bei dem ein gegebenes Voxel bei $p = 0.001$ signifikant war. Dieser letzte Wert stammt offenbar vom Signifikanzniveau $\alpha = 0.001$ des einseitigen Einstichproben-t-Test, mit dem

⁹ Das HCP ist ein 2009 gestartetes, von den National Institutes of Health gefördertes Projekt mit dem Ziel, eine Karte des menschlichen Nervensystems anzulegen. Beteiligt sind unter anderem die Washington University und das Massachusetts General Hospital, Lehrkrankenhaus für die Universität Harvard. Projektwebseite: humanconnectomeproject.org [abgerufen am 06.10.2019].

¹⁰ $N = 481$ Probanden umfasste das Arbeitsgedächtnis-Experiment, $N = 479$ das Motorik-Experiment, $N = 471$ das Emotionen-Experiment und $N = 479$ das Glücksspiel-Experiment.

die Autoren die Leistungsfähigkeit statistisch berechneten. Die Ergebnisse für die Aufgaben zu Arbeitsgedächtnis, Motorik und Emotionsverarbeitung ähnelten sich darin, dass der größte Anstieg an Leistungsfähigkeit zwischen einer Stichprobengröße von $N = 20$ und $N = 40$ verzeichnet wurde. Für $N > 40$ stieg die Leistungsfähigkeit stetig an. Eine Leistungsfähigkeit von 50% wurde für manche Hirnregionen bereits zwischen $N = 10$ und $N = 20$ errechnet. Mit steigender Stichprobengröße erhöhte sich auch die Zahl über den Schwellenwert hinaus aktivierter Voxel, was bedeutet, dass Aktivierungen in manchen Hirnregionen erst bei höheren Stichprobengrößen erfasst wurden. Für manche Hirnregionen wurde eine Leistungsfähigkeit von 90% bei Stichprobengrößen zwischen $N = 60$ und $N = 100$ erreicht. Das gilt besonders für das Arbeitsgedächtnis-Experiment. Die Ergebnisse für die Studie mit der Glücksspielaufgabe weichen von denen der übrigen Aufgabentypen ab. Hier sind viel größere Stichprobengrößen nötig, um eine für die Autoren akzeptable Leistungsfähigkeit zu erreichen. Manche Hirnregionen erreichen 80% Leistungsfähigkeit mit einer Stichprobengröße von $N > 20$, während andere, wie das ventrale Striatum, bei einer Größe von $N > 40$ eine Leistungsfähigkeit von 50% erreichen. Erst bei $N > 100$ werden 80% und mehr Leistungsfähigkeit erreicht. Für Neuroökonomien sollten diese Ergebnisse besonders interessant sein, da sich ihre Experimentaufgaben zumeist um Spiele drehen. Geuter et al. Relativieren diese hohen Zahlen jedoch mit der Annahme, dass, da gerade das ventrale Striatum als Schlüsselregion für *Value Learning* angesehen wird, die niedrigen Werte für Leistungsfähigkeit wahrscheinlich mit dem im ursprünglichen Glücksspiel-Experiment verwendeten Block-Design zusammenhängen. Warum oder inwiefern sie bei beispielsweise einem *Event-related-Design* andere Ergebnisse erwarten würden, sagen die Autoren leider nicht. Generell schließen sie, dass für individuelle fMRT-Studien eine Stichprobengröße von $N = 40$ empfehlenswert ist, wenn man sowohl die Reliabilität als auch die Kosten betrachtet. Das sind deutlich mehr Probanden, als in den in dieser Arbeit betrachteten Studien rekrutiert wurden. Neuroökonomien sollten diese Kritik also ernst nehmen. Dafür hilft eine konsequente Kontrolle der Publikationen nicht nur durch Peer-Reviewer, sondern auch durch die wissenschaftliche Gemeinschaft, die schlechte Forschungsqualität nicht zum Wohle spannender Ergebnisse unter den Tisch fallen lassen darf. Immerhin lässt sich sagen, dass die Neurowissenschaften dieses Problem offenbar inzwischen selbst sehen und Studien in diese Richtung betreiben, um klare statistische Richtwerte zu erhalten. Umsetzen, allerdings, müssen sie die guten neuen Vorsätze auch.

6.1.4 Schlechtes Forschungsdatenmanagement

Diese Kritik dreht sich um die Veröffentlichungspolitik der Neurowissenschaftler bezüglich der Rohdaten ihrer Experimente. Ökonomen beklagen, dass Neurowissenschaftler ihre Messdaten nicht preisgeben und so die Nachvollziehbarkeit ihrer Experimente verringern.

Forscher wie Ortmann (Ortmann 2008) und Harrison (Harrison 2008a) sehen es als skandalös, dass Neuroökonomien ihre Rohdaten nicht veröffentlichen, auch nicht auf Nachfrage. Harrison (2008a, 315) berichtet seine Erfahrungen:

„A serious issue arises, however, from one of the dark secrets of this field: no neural data is apparently ever made public. For some years requests were made to prominent authors in neuroeconomics, particularly economists with significant and well-earned reputations in experimental economics. Not one has provided data. In some cases the economist could not get the neuroscience co-author to release data, perhaps reflecting traditions in that field. [...] In many cases the economist in question gladly provided any ‘behavioural’ non-image data, although there are several prominent publications for which even those data are being withheld ‘pending further analysis and publications’.“

Dabei sieht Harrison in dieser Intransparenz ein größeres Problem, das die Vertrauenswürdigkeit der gesamten Neuroökonomie gefährdet. Denn ohne Datenveröffentlichung keine Reproduktionsmöglichkeit und ohne Reproduktionsmöglichkeit keine Nachprüfbarkeit der Daten und Schlüsse. Daher kann auch kein Vertrauen in Daten, Wissenschaftler, ja die gesamte Neuroökonomie entstehen (Harrison 2008b, 539):

„Maybe we need to spell out why data sharing is such an issue. Trust in the results of a field arises through replication of inferences, and without sharing of data there can simply be no replication. And here we mean ‘replication of inferences under alternative structural and stochastic statistical assumptions’, not replication in the sense of one neuroeconomist thinking he gets the same finding as another. This is, after all, the problem with data privacy: the rest of us are unable to critically examine the importance of statistical assumptions in driving results.“

Auch die Antwort von Neuroökonomien auf diese Kritik findet er nicht hilfreich (Harrison 2008b, 539):

„Camerer (2008[b]) does not deny the data privacy problem, but still asks us to trust that results and interpretations wouldn’t change if we looked at the data ourselves. Unacceptable.“

Nicht nur aus Harrisons Sicht ist dieser Vorschlag tatsächlich inakzeptabel. Allein auf die korrekte Arbeit der Wissenschaftler zu vertrauen, ist keine Lösung, zumal für jemanden, der ihren Schlussfolgerungen nicht traut (siehe unten). Zudem werden Replikationen ja nicht in erster Linie dazu genutzt, die Vertrauenswürdigkeit der Wissenschaftler zu überprüfen, sondern experimentelle Phänomene zu reproduzieren und Auswirkungen kleinerer oder größerer Änderungen am Experimentdesign zu untersuchen. Das ist ein wichtiges Hilfsmittel, um die Experimentergebnisse und Schlussfolgerungen zu stützen oder falsifizieren. Ganz klar sind solche Vorhaben jedoch hinfällig ohne Vergleichsdaten und das könnte, da hat Harrison recht, ein Problem für die Neuroökonomie werden.

Es ist allerdings auch nicht der Fall, dass gar keine Daten veröffentlicht werden. Denn es werden durchaus Daten veröffentlicht, heute mehr als noch vor ein paar Jahren, was sicher mit der steigenden Zahl an Wissenschaftsförderern zu tun, hat, die Open-Data-Mandate in ihre Förderrichtlinien integrieren, wie beispielsweise die EU-Kommission. Schon länger relativ detailliert veröffentlicht werden die vorgenommenen Schritte der Datenvorverarbeitung, das haben wir auch in Kapitel 5 gesehen. Diese sind genauso wichtig wie die Messdaten selbst, nicht nur, weil nur so nachvollziehbar wird, was mit den Daten passiert ist und ob sie nach den geltenden, akzeptierten Standards verarbeitet wurden, sondern auch, weil gerade die Datenvorverarbeitung, wie wir gesehen haben, Anlass zu Kritik über die Zuverlässigkeit der Analysen gibt. Ob sie allerdings ausreichen für potenzielle Replikationsstudien, bei denen die Daten neu analysiert oder ganz neu erhoben und verglichen werden können, ist eine andere Frage. Hierfür sind die genauen Messergebnisse des Experiments interessant, um eigene Messungen vergleichen zu können, oder auch mit demselben Datensatz neue Analysen zu neuen Fragestellungen durchzuführen, wie Geuter et al. in ihrer Studie (Geuter et al. 2018, s. o. Abschn. 6.1.3). Replikationen sind in etablierten Experimentalwissenschaften Normalität und ein probates Mittel, um Hypothesen und Theorien weiter auszutesten. Daher wäre eine Null-Publikations-Politik hinderlich für die Neuroökonomie auf dem Weg zur anerkannten Wissenschaft. Allerdings müssen für Datenpublikationen die rechtlichen Rahmenbedingungen stimmen. Bei neuroökonomischen Studien fallen sensible persönliche Daten an, die unbedingt ausreichend anonymisiert werden müssen, bevor sie veröffentlicht werden können. Außerdem müssen Vorgaben von Geldgebern bedacht werden. Nicht alle Förderer erlauben oder gar fordern eine Veröffentlichung von Daten. Gerade bei privatwirtschaftlichen Geldgebern sollen aus Studienergebnissen für gewöhnlich vermarktbar Produkte entstehen, was durch die Veröffentlichung von Forschungsdaten gefährdet werden könnte. Vermutlich will auch kein privater Geldgeber die kostenintensiv generierten Daten veröffentlicht wissen, bevor

sie nicht auf eine mögliche Verwertung für weitere Forschungszwecke analysiert worden sind. Vielleicht lassen sich mit einem Datensatz auch noch andere Forschungsfragen beantworten und Publikationen generieren. Daher ist es nicht verwunderlich, wenn Harrison unter vielen Artikeln den Hinweis findet, dass die Daten nicht veröffentlicht werden mit Rücksicht auf mögliche zukünftige Publikationen. Damit will ich sagen: Es ist nicht ausgeschlossen, dass einige Forschungsteams ihre Daten veröffentlichen würden, es aber nicht dürfen.

Dieses gesamte Themengebiet um Veröffentlichung von Forschungsdaten ist in jüngster Zeit stark in Bewegung gekommen. Große Forschungsförderer wie die Deutsche Forschungsgemeinschaft DFG fordern in ihren Leitlinien zum Umgang mit Forschungsdaten, diese „in der eigenen Einrichtung oder in einer fachlich einschlägigen, überregionalen Infrastruktur“ (DFG 2015, 1) für mindestens 10 Jahre zu archivieren. Damit sind sowohl ganze Einrichtungen als auch einzelne Forscher gezwungen, sich über ihr Forschungsdatenmanagement Gedanken zu machen und die bisher nicht selten anzutreffende Praxis der Selbstarchivierung auf USB-Sticks in der Schreibtischschublade zu ersetzen. Die Zahl der Institutionen und Fachbereiche, die solche fachspezifischen Datenrepositorien anbieten, wächst derzeit rasch, was immer mehr Forschern aus immer mehr Fachgebieten die Möglichkeit gibt, ihre Daten professionell zu archivieren, um sie auch in zehn Jahren noch abrufen zu können. Das ist eine Voraussetzung dafür, die Daten in einem zweiten Schritt zur Nachnutzung durch dritte Forschungseinrichtungen zur Verfügung zu stellen. Dazu bedarf es dann nicht nur fachkundiger Rechenzentren mit großen Speicherkapazitäten. Momentan gibt es vereinzelt Wissenschaftler, die ihre Forschungsdaten aus Experimenten selbst auf ihrer eigenen Webseite veröffentlichen. Das birgt natürlich nicht nur die Frage nach der Auffindbarkeit der Daten, vor allem rechtliche Fragen sind derzeit ein Problem, gerade bei personenbezogenen oder personenbezieharen Erhebungen. Denn niemand möchte Patientendaten in Repositorien zur freien Verfügbarkeit für jeden interessierten Forscher sehen. Besonders DNA-Sequenzen sind hier ein stark diskutiertes Thema. Solche Probleme können auch Daten aus neurowissenschaftlichen Experimenten betreffen. Kurzum: In der Problematik mit nicht zugänglichen, nicht überprüfbaren Experimentdaten sind die Gegebenheiten gerade sehr in Bewegung, die Forschungslandschaft verändert sich stark hin zu mehr Transparenz und Zugänglichkeit. Es sind allerdings noch rechtliche Fragen zu klären, bevor sich jeder Interessierte die Rohdaten von neuroökonomischen Experimenten herunterladen und selbst beurteilen kann.

Harrison (Harrison 2008b, 539) beklagte auch, dass es durch eine Geheimhaltung der Daten unmöglich wird, zu beurteilen, welchen Wert die statistischen Annahmen haben, die Schlussfolgerungen zugrunde liegen („[...] the rest of

us are unable to critically examine the importance of statistical assumptions in driving results.“). Das mag schon stimmen. Wenn man als Leser dem statistischen/logischen Vermögen oder der wissenschaftlichen Integrität des Forschungsteams nicht traut, dann scheinen Messdaten eine relevante Informationsquelle. Diese Überlegung übersieht aber die Tatsache, dass, wie wir in Abschnitt 5.2, S. 115–118, gesehen haben, die statistische Auswertung der Messdaten heute automatisiert von Softwarespezialistischer internationaler Anbieter durchgeführt wird. Die Versiertheit des Forschungsteams im Berechnen statistischer Größen hat also keinen sehr großen Einfluss auf die statistische Auswertung der Rohdaten, denn welche Werte wie bewertet werden, ist in der Software festgelegt. Das Forschungsteam sollte daher immerhin eine für ihre Zwecke passende Software auswählen. Für welche sie sich entschieden haben, ist normalerweise relativ detailliert in den zur Studie gehörenden Publikationen beschrieben (siehe beispielsweise Zaki et al. 2011) und kann daher von allen Lesern eingesehen und beurteilt werden.

6.1.5 *Reverse Inference* als gängige Praxis?

Neben den Bedenken, Neurowissenschaftler könnten ihre Daten nicht korrekt bewerten, gibt es Kritiken, die Neurowissenschaften pflegten eine Kultur der logisch falschen Schlüsse.

Der kognitive Psychologe Russell Poldrack weist in einem Artikel auf ein Problem hin, das mit der Praxis der *reverse Inference* einhergeht und von den Nutzern bildgebender Verfahren stärker beachtet werden sollte.

Wie *reverse Inference* in bildgebenden Studien funktioniert, erklärt Poldrack (Poldrack 2006, 59) so:

„Functional neuroimaging techniques such as functional magnetic resonance imaging (fMRI) provide a measure of local brain activity in response to cognitive tasks undertaken during scanning. These data allow the cognitive neuroscientist to infer something about the role of particular brain regions in cognitive function. However, there is increasing use of neuroimaging data to make the opposite inference; that is, to infer the engagement of particular cognitive functions based on activation in particular brain regions. [...]

The usual kind of inference that is drawn from neuroimaging data is of the form ‘if cognitive process *X* is engaged, then brain area *Z* is active’. Perusal of the discussion sections of a few fMRI articles will quickly reveal, however, an epidemic of reasoning taking the following form:

- (1) In the present study, when task comparison A was presented, brain area Z was active.
- (2) In other studies, when cognitive process X was putatively engaged, then brain area Z was active.
- (3) Thus, the activity of area Z in the present study demonstrates engagement of cognitive process X by task comparison A.

This is a ‘reverse inference’, in that it reasons backwards from the presence of brain activation to the engagement of a particular cognitive function.“

In vielen Studien sieht Poldrack die *reverse Inference* informell verwendet, um das Auftreten unerwarteter Aktivierungen in einer bestimmten Hirnregion mit Verweis auf andere Studien zu erklären, in denen eine Aktivierung derselben Region aufgetreten war. Allerdings sieht er auch viele Studien, in denen die *reverse Inference* ein zentrales Merkmal ist. Auf die eine oder andere Weise verwendet beinahe jede bildgebende Studie *reverse Inference*. Poldrack weist darauf hin, dass diese Form der Schlussfolgerung deduktiv nicht gültig ist und tatsächlich einen logischen Fehlschluss begeht. Das obenstehende Schlussystem wäre deduktiv gültig, würde (2) aussagen, dass Hirnregion Z dann und nur dann aktiv war, wenn der kognitive Prozess X beteiligt war. Davon kann aber nicht ausgegangen werden.

Neben diesem logischen Problem sieht Poldrack (Poldrack 2006, 60, 63) die *reverse Inference* aber auch als ein nützliches Hilfsmittel bei der Entdeckung neuer Fakten über die den kognitiven Prozessen zugrundeliegenden neuronalen Mechanismen. *Reverse Inference* könnte neue Hypothesen anregen, die dann in weiteren Experimenten getestet werden könnten. Allerdings sollte sie mit Vorsicht angewandt werden, wie eine Veranschaulichung anhand des Satzes von Bayes zeigt (Poldrack 2006, 60):

$$\begin{aligned}
 & \text{„}P(\text{COG}_X|\text{ACT}_Z) \\
 & \quad P(\text{ACT}_Z|\text{COG}_X) P(\text{COG}_X) \\
 & = \text{-----} \\
 & \quad P(\text{ACT}_Z|\text{COG}_X) P(\text{COG}_X) + P(\text{ACT}_Z| \neg\text{COG}_X) P(\neg\text{COG}_X)
 \end{aligned}$$

where COG_X refers to the engagement of cognitive process X and ACT_Z refers to activation in region Z. (It should be noted that the prior $P(COG_X)$ is always conditioned on the particular task being used, and should more properly be termed $P(COG_X|TASK_y)$; however, for the purposes of simplicity I have omitted this additional conditionalization.)“

Die Wahrscheinlichkeit, dass eine mit *reverse Inference* nach oben zitiertem Schema (1) bis (3) geschlussfolgerte kognitive Funktion X vorliegt, wird hier also als bedingte Wahrscheinlichkeit berechnet: Die Wahrscheinlichkeit, dass der kognitive Prozess X vorliegt, gegeben eine Aktivierung in Hirnregion Z. Diese berechnet sich nach Bayes als Division mit dem Produkt aus der bedingten Wahrscheinlichkeit für eine Aktivierung in Hirnregion Z gegeben der kognitive Prozess X, und der Anfangswahrscheinlichkeit für das Auftreten des kognitiven Prozesses X als Dividend. Der Divisor ist die Summe aus der bedingten Wahrscheinlichkeit für eine Aktivierung in Hirnregion Z gegeben kognitiver Prozess X multipliziert mit der Anfangswahrscheinlichkeit für das Auftreten von X, und der bedingten Wahrscheinlichkeit für eine Aktivierung in Hirnregion Z gegeben das Nicht-Vorhandensein des kognitiven Prozesses X multipliziert mit der Anfangswahrscheinlichkeit für das Nicht-Auftreten von X. Der Divisor kann auch zusammengefasst werden zur Basiswahrscheinlichkeit $P(ACT_Z)$, also der Wahrscheinlichkeit für das Auftreten einer Aktivierung in Hirnregion Z.

Poldrack verdeutlicht mit dieser Aufstellung, dass aus $COG_X \rightarrow ACT_Z$ nicht einfach $ACT_Z \rightarrow COG_X$ gefolgert werden kann. Um die *reverse Inference* dennoch als Hilfsmittel einsetzen zu können, schlägt Poldrack also vor, sich nach der Wahrscheinlichkeit zu richten, mit der die *reverse Inference* zutrifft, also mit der bedingten Wahrscheinlichkeit für den kognitiven Prozess X gegeben die Aktivierung in Hirnregion Z. Nach Poldrack hängt der Grad der Überzeugung in einer *reverse Inference* sowohl von der Selektivität der neuronalen Antwort (d. h. dem Verhältnis einer prozess-spezifischen Aktivierung zu der allgemeinen Wahrscheinlichkeit einer Aktivierung dieser Hirnregion über alle Aufgaben) ab, als auch von der A-priori-Überzeugung in die Beteiligung des kognitiven Prozesses X gegeben die Manipulation durch die Experimentaufgabe [$P(COG_X)$].

Die Darstellung in Bayes Formel vereinfacht Poldrack die Aufstellung zweier Möglichkeiten, wie das Vertrauen in die *reverse Inference* verstärkt werden kann (Poldrack 2006, 62–63): Erhöhe die Selektivität der neuronalen Antwort der interessierenden Hirnregion, oder erhöhe die A-priori-Wahrscheinlichkeit des fraglichen kognitiven Prozesses. Die Selektivität liegt außerhalb der Kontrolle der Forschungsteams, obwohl Poldrack eine Schätzung immerhin für möglich

hält. Auch die Analyse eines Sets an Hirnregionen, die als Netzwerk funktionieren, könnte eine größere Selektivität bieten, ebenso wie eine kleiner gewählte interessierende Hirnregion mehr Zutrauen bieten könnte.

Die A-priori-Wahrscheinlichkeit für die Beteiligung der interessierenden Hirnregion dagegen unterliegt zumindest zu einem gewissen Grad der Kontrolle der Forschungsteams. Sie können oft Experimentaufgaben so auswählen, dass diese Wahrscheinlichkeit maximiert wird. So wird auch die Wahrscheinlichkeit für das Zutreffen der *reverse Inference* erhöht. Poldrack schränkt aber ein, dass diese Strategie eher für Studien anwendbar ist, die darauf zielen eine spezifische *reverse Inference* zu erreichen, als für solche, bei denen die *reverse Inference* verwendet wird, um post hoc eine Erklärung für ein bestimmtes Ergebnis zu liefern.

Grundsätzlich ist Poldrack ein Befürworter der Verwendung von *reverse Inference* in bildgebenden kognitiven Studien. Er hält sie für nützlich, um neue Hypothesen aufzustellen, denen in weiteren Studien nachgegangen werden kann. Trotzdem warnt er beim Umgang mit *reverse Inference* zur Vorsicht, vor allem dort, wo die A-priori-Wahrscheinlichkeit für die Beteiligung eines kognitiven Prozesses und die Aktivierungs-Selektivität einer Hirnregion niedrig sind. Dann ist die Gefahr, aus einer Abduktion einen Fehlschluss zu machen, besonders hoch.

Harrison (2008b, 536–538) seinerseits sieht ebenfalls den Wert und die Nützlichkeit, den solches Schließen für das Fortkommen der Neurowissenschaften beziehungsweise der Neuroökonomie hat, hält diese Praxis der Wissensgenerierung über formelle Fehlschlüsse aber für keine gute Basis für die Neuroökonomie. Er befürchtet, dass die Praxis der *reverse Inference* in den einzelnen Fällen als harmlos angesehen wird, aber im Ganzen dann kein stabiler Wissensfortschritt entsteht, wie es die Neuroökonomien planen, sondern der Verlust an Logik sich aufsummiert, sodass die Daten generell an Verlässlichkeit einbüßen. Er sieht zwar eine Aufmerksamkeit für und Auseinandersetzung mit diesem Problem in den kognitiven Neurowissenschaften, in der neuroökonomischen Literatur findet er diese Problematik jedoch oft beschönigt, statt beachtet. Hier sollten die Neuroökonomien von den Neurowissenschaftlern lernen und dieselbe Vorsicht walten lassen, die Poldrack anmahnt.

An dieser Stelle schließe ich mich Harrison an. Natürlich ist es keine gute Idee, *reverse Inference* als allgemein gültige Schlussmöglichkeit einzusetzen und die Ergebnisse als Grundlage, nicht Ausgangspunkt, weiterer Forschungen zu verwenden. Gerade die Neuroökonomien müssen die wissenschaftlichen Praktiken mehrerer Mutterwissenschaften zusammenbringen, sodass auch andere Forscher aus den Mutterwissenschaften überzeugt werden können. Ich stimme mit Harrison überein, dass die Anwendung von *reverse Inference* in den jeweils einzelnen Studien als nicht Besorgnis erregend erscheinen mag, der Gesamtschaden aber

erheblich sein kann. Poldrack ist also auf dem richtigen Weg, wenn er *reverse Inference* nicht als ausreichende Erkenntnisquelle für Post-hoc-Erklärungen von Aktivierungen sieht, sondern als Hilfsmittel für das Aufstellen und Überprüfen neuer Hypothesen. Dafür liefert er ein gutes theoretisches Hilfsmittel mit seiner Formulierung des Satzes von Bayes. Sie kann gut veranschaulichen, wie groß die Wahrscheinlichkeit für das Zutreffen der Schlussfolgerung aus einer *reverse Inference* in etwa ist und bietet damit eine Grundlage zur Einschätzung, ob man über sie zu neuer, wenn auch probabilistischer, Evidenz gelangen kann. Jeder Forscher muss sich über die Instabilität von *reverse Inference* im Klaren sein und bei Schlüssen aus *reverse Inference* in Publikationen klar darauf hinweisen, dass sie auf *reverse Inference* beruhen.

6.2 Die systematischen Kritiken

Nach den prozeduralen werden wir nun die systematischen Kritiken betrachten, mit denen sich die Neuroökonomie auseinandersetzen muss. Sie sind von substantiellerer Art und stellen alle mehr oder weniger gravierende Probleme für die Zusammenarbeit von Verhaltensökonomern, kognitiven Neurowissenschaftlern und Psychologen dar. Im Gegensatz zu den Kritiken, die wir gerade betrachtet haben, betreffen sie nicht nur die tägliche Arbeit der Neurowissenschaftler und Ökonomen. Ihre Wurzeln liegen vielmehr bereits in der Auffassung der verschiedenen Wissenschaftler darüber, wie ihre Wissenschaft funktioniert, beziehungsweise funktionieren soll. Mit ihren Gedanken darüber haben sich vor allem die Ökonomen Gul und Pesendorfer hervorgetan, die zum Teil harsche Kritik an der Neuroökonomie üben. Die im Folgenden besprochenen Kritiken stammen daher aus ihrem Artikel „The Case for Mindless Economics“ aus dem Sammelband „The Foundations of Positive and Normative Economics“ (Caplin et al. 2008).

6.2.1 Ökonomen sind keine Therapeuten

Generell kritisieren Gul und Pesendorfer die unterschiedliche Herangehensweise von Ökonomen und Neuroökonomern, wenn es um entscheidungstheoretische Fragen, Begriffe oder Konzeptionen geht. Als Beispiel zeigen sie eine neuroökonomische und eine ökonomische Definition von Risikoaversion (Gul et al. 2008, 5):

„Much aversion to risks is driven by immediate fear responses, which are largely traceable to a small area of the brain called the amygdala.’ [Camerer, Loewenstein, Prelec, 2004: 567]

‘A decision maker is (globally) risk averse, ... if and only if his von Neumann-Morgenstern utility is concave at the relevant (all) wealth levels.’ [Ingersoll, 1987: 37]“

Gul und Pesendorfer machen darauf aufmerksam, dass beide Definitionen unterschiedliches, aber spezialisiertes Vokabular verwenden: Für Ingersoll ist Risikoaversion eine Einstellung gegenüber Lotterien, während Risikoaversion für Camerer, Loewenstein und Prelec ein viel breiterer Begriff zu sein scheint, der sich ebenso auf Entscheidungen anwenden ließe, bei denen es um Flugreisen geht. Gul und Pesendorfer betonen, dass es keinen Sinn macht, zu fragen, welche Definition das bessere Verständnis von Risikoaversion liefert, oder welche zutreffender sei. Es ist in ihren Augen wenig sinnvoll, darauf zu beharren, die ökonomische Auffassung von Risikoaversion sei falsch, während die psychologische richtig sei. Sie sind aber der Meinung, dass Neuroökonomien genau das tun. Das beziehen sie vor allem auf die neuroökonomische Kritik an der ökonomischen Wohlfahrtsanalyse, die Gul und Pesendorfer für charakteristisch für die Neuroökonomie halten. Für sie ist Neuroökonomie definiert als (Gul et al. 2008, 3):

„We identify neuroeconomics as research that implicitly or explicitly makes either of the following two claims:

Assertion 1: Psychological and physiological evidence (e.g., descriptions of hedonic states and brain processes) is directly relevant to economic theories. In particular, it can be used to support or reject economic models or even economic methodology.

Assertion 2: What makes individuals happy (“true utility”) differs from what they choose. Economic welfare analysis should use true utility rather than the utilities governing choice (“choice utility”).“

Mit Gul und Pesendorfers Kritik an *Assertion 1* werden wir uns weiter unten noch beschäftigen, an dieser Stelle geht es vor allem um *Assertion 2*, die die neuroökonomische Einstellung zur ökonomischen Wohlfahrtsanalyse beschreibt. Sie widerspricht deutlich dem *Revealed-Preference-Modell*, einem der Grundpfeiler der neoklassischen Ökonomie. Es ist nicht verwunderlich, dass Gul und Pesendorfer das nicht begrüßen. Ihrer Meinung nach (Gul et al. 2008, 5) missverstehen die Neuroökonomien die ökonomische Definition von Wohlfahrt als eine Glückseligkeitstheorie. Und als eine solche Glückseligkeitstheorie macht sie in den Augen der Neuroökonomien einen schlechten Job. Daher wollen sie gegen diese gängige

ökonomische Definition der Wohlfahrtstheorie Beweise finden. Gul und Pesendorfer machen aber klar, dass die ökonomische Standarddefinition von Wohlfahrt adäquat ist, da die klassische Ökonomie keine, wie Gul und Pesendorfer es nennen, therapeutischen Ambitionen hat. Sie versucht ihrer Meinung nach nicht, die Entscheidenden zu verbessern, sondern versucht, zu bewerten, wie ökonomische Institutionen (wie Organisationsstrukturen oder Handelsmechanismen) das Verhalten von Agenten herbeiführen, gleichgültig ob dieses Verhalten psychologisch gesund ist oder nicht. Daher ist die ökonomische Wohlfahrtstheorie keine Theorie darüber, wie die Menschen Glückseligkeit erreichen können. Neuroökonomien aber erwarten offenbar, dass jüngste Entwicklungen in Psychologie und Neurowissenschaften alte philosophische Fragen danach, was Glück ist, beantworten könnten, wobei aus diesen Antworten eine neue Auffassung von Wohlfahrt entstehen soll. Damit, meinen Gul und Pesendorfer, wäre vielleicht Therapeuten oder Medizinern dabei geholfen, das Glück ihrer Patienten zu erhöhen, aber auf die Ökonomie sind sie nicht anwendbar. Die Neuroökonomien missverstehen in ihren Augen nicht nur die Wohlfahrtskonzeption der Ökonomen, sondern auch ihr Verhältnis zum ökonomischen Agenten (Gul et al. 2008, 6):

„Neuroeconomic welfare analysis assumes a relationship between the economist and economic agents similar to the therapist-patient relationship. Normative economics is therefore identified with effective therapy. [...] Neuroeconomists seek a welfare criterion that is appropriate for an economist who is part social scientist and part advocate/therapist, someone who not only analyzes economic phenomena but also plays a role in shaping them.“

Sie nehmen also an, Neuroökonomien sehen Ökonomen als Helfer oder Betreuer bei der Steigerung der Wohlfahrt ökonomischer Agenten. Diese Rolle, oder auch nur die Intention, diese Rolle einzunehmen, weisen Gul und Pesendorfer aber weit von sich und allen Ökonomen. Sie sehen in der neuroökonomischen Verwerfung der ökonomischen Wohlfahrtstheorie sogar Tendenzen zur Paternalisierung der Beziehung von ökonomischen Institutionen zu Individuen. Das machen sie an einem Zitat von Kahneman fest (Gul et al. 2008, 6):

„For example, Kahneman [1994: 20] suggests that there is ‘a case in favour of some paternalistic interventions, when it is plausible that the state knows more about an individual’s future tastes than the individual knows presently.’ Hence, the goal of welfare economics and perhaps the goal of all economics is to affect changes that result in greater happiness to all.“

Um dieses Ziel zu erreichen, sehen Gul und Pesendorfer die NeuroökonomInnen bereit, die Hilfe des Staates einzuholen, der gegebenenfalls Fakten verbirgt und im Namen der zukünftigen Ichs der Individuen Entscheidungen trifft. Mit solchen Methoden, und der Rolle, in die sie die Ökonomie durch die Neuroökonomie gedrängt sehen, hat die tatsächliche Ökonomie ihrer Meinung nach nichts zu tun. In ihren Augen ist das Standard-Wohlfahrtskriterium der Ökonomen nicht dazu gedacht, die Befürwortung therapeutischer Interventionen zu erleichtern. Sie erläutern, dass die ökonomische Herangehensweise von einer Trennung der Rolle des Ökonomen als Sozialwissenschaftler von der Rolle als Berater oder Fürsprecher ausgeht, die manche Ökonomen einnehmen. Diese Trennung sehen sie als wertvoll, weil sie den Ökonomen ermöglicht, verschiedene Institutionen zu analysieren und zu vergleichen, ohne dass sie sich über schwierige philosophische Fragen (wie die oben erwähnte, danach, was Glück ist) einig sein müssen.

Gul und Pesendorfer kritisieren hier also zwei Dinge: Sie sehen bei den NeuroökonomInnen a) ein Missverständnis von Ökonomie als aktive, verändernde Therapie, und b) den Wunsch nach Paternalisierung der Beziehung von ökonomischen Institutionen zu Individuen. Für Gul und Pesendorfer sind Ökonomen nicht halb Sozialwissenschaftler, halb Therapeut, sie analysieren ökonomische Phänomene und formen sie nicht mit.

Die Kritik in a) kann man gelten lassen. An Universitäten und ähnlichen Einrichtungen forschende Ökonomen sehen es grundsätzlich nicht als ihre Aufgabe, mithilfe ihrer Wohlfahrtsanalyse ökonomischen Individuen in therapeutischer Weise dabei zu helfen, ihr Glück zu erhöhen. Die Rede von Therapie, die Gul und Pesendorfer hier ins Spiel bringen, ist allerdings ein bisschen polemisch. Zudem darf nicht vergessen werden, dass diese Kritik von Gul und Pesendorfer nur dann zutrifft, wenn die oben genannte *Assertion 2* über die Neuroökonomie korrekt ist. Die stammt allerdings allein aus der Feder von Gul und Pesendorfer und entspricht lediglich ihrer Einschätzung der Neuroökonomie. Die Aussagen in *Assertion 2* haben Gul und Pesendorfer nicht mit einem Text von NeuroökonomInnen belegt. Sie schaffen sich also die Grundlage für ihr eigenes Argument selbst, die dadurch nicht viel wert ist.

In Fall von Kritikpunkt b) legen Gul und Pesendorfer eine Belegstelle von Kahneman vor, in der er für die Anwendung paternalistischer Interventionen eintritt. Für solche Zwecke sehen die Autoren die Wohlfahrtskonzeption der Ökonomie nicht gedacht und nicht geeignet. Das ist nach ihrer Ansicht weder die Aufgabe noch die Kompetenz der Ökonomen und ihrer Wohlfahrtsanalyse. Allerdings frage ich mich an dieser Stelle, warum Gul und Pesendorfer dieses Kahneman-Zitat bringen. Es ist mir nicht bekannt, dass sich Kahneman als besonderer Verfechter der Neuroökonomie hervorgetan hätte, vor allem nicht 1994, als

die Neuroökonomie noch höchstens ein entfernter Traum war. Sein Zitat steht auch in keinem zwingenden Zusammenhang mit neuroökonomischen Ideen. Was also haben sich Gul und Pesendorfer an dieser Stelle gedacht? Mir drängt sich der Eindruck auf, die Autoren bauen hier, beabsichtigt oder unbewusst, eine (extreme) Position zum Strohmann auf und bekämpfen ihn. Das beeindruckt vielleicht manche Leser, ist aber tatsächlich keine belastbare Kritik gegenüber der Neuroökonomie.

6.2.2 Neuroökonomie ist irrelevant für Ökonomie

In der folgenden Kritik geht es um Guls und Pesendorfers Argumentation für eine grundsätzliche Unvereinbarkeit von Neuroökonomie und Ökonomie, die sie auf die unterschiedliche Beschaffenheit der beiden Wissenschaften zurückführen. Sie halten die Unterschiede für so gravierend, dass sie unüberwindbar sind. Ihre Argumentation ist gegen die *Assertion 1* gerichtet, die Gul und Pesendorfer der Neuroökonomie zuschreiben (siehe oben Abschn. 6.2.1, S. 152).

Gul und Pesendorfer stimmen nicht mit der darin beschriebenen Idee überein. Für sie ist klar, dass psychologische und physiologische Erkenntnisse aus den Neurowissenschaften oder der Psychologie keine direkte Relevanz für die Ökonomie haben und auch nicht dazu verwendet werden können, ökonomische Modelle oder Methoden zu stützen oder zu verwerfen. Gul und Pesendorfer halten die Neuroökonomie und ihre Ideen und Erkenntnisse vielmehr für irrelevant für die Ökonomie. Das machen sie an zwei Gründen fest: Erstens, die Neuroökonomie spricht keine ökonomischen Probleme an, und zweitens, die Ökonomie spricht keine neurowissenschaftlichen Probleme an. Der zweite Punkt ist vielleicht noch etwas interessanter als der erste, aber zusammen ergeben sie ein Bild von der Neuroökonomie und auch der Ökonomie, wie Gul und Pesendorfer sie sehen.

Beginnen wir mit dem ersten Aspekt: Die Neuroökonomie spricht die ökonomischen Probleme nicht an. Gul und Pesendorfer (Gul et al. 2008, 9–11) stellen klar, dass die Neuroökonomie und die Ökonomie unterschiedlichen Fragestellungen nachgehen, unterschiedliche Ziele verfolgen, dabei unterschiedliche Mittel anwenden und mit unterschiedlichen Quellen und Arten von Daten arbeiten. So sehen sie als zentrale Fragen der Neuroökonomie: Wie treffen Individuen ihre Wahlen? Wie effektiv sind sie darin, diejenigen Entscheidungen zu treffen, die ihr eigenes Wohlergehen steigern? Dagegen ist das zentrale Anliegen der Ökonomen ihrer Meinung nach, zu analysieren, wie die Entscheidungen verschiedener Individuen mit verschiedenen gegebenen Zielen innerhalb einer bestimmten

institutionellen Umgebung interagieren. Diese unterschiedlichen Zielsetzungen verlangen laut Gul und Pesendorfer nach unterschiedlichen Herangehensweisen, Begriffsdefinitionen, Daten und Hilfsmitteln, die deshalb so unterschiedlich sind, weil sie sich in der jeweiligen Weise für die jeweilige Disziplin als nützlich herausgestellt haben. Gul und Pesendorfer sind aber der Meinung, dass Neuroökonomien das nicht einsehen oder verstehen. Ihrer Ansicht nach sehen die Neuroökonomien bei Psychologie, Neurowissenschaften und Ökonomie die gleichen Zielsetzungen. Als Beispiel dient ihnen ein Zitat aus einem Aufsatz von Camerer, Loewenstein und Prelec (Camerer et al. 2004, 572–573), in dem die Autoren von Studenten berichten, die sich erstaunt darüber zeigen, dass die gleichen Begriffe, beispielsweise “trust”, in unterschiedlichen Disziplinen unterschiedlich definiert sind. Sie erläutern darauf:

„It is possible that a biological basis for behavior in neuroscience, perhaps combined with all-purpose tools like learning models or game theory, could provide some unification across social sciences.“

Gul und Pesendorfer (Gul et al. 2008, 10–11) betonen, dass Ökonomie und Psychologie keine konkurrierenden Modelle liefern, man sie also auch kaum miteinander vergleichen kann. Die Spieltheorie ist für sie nicht, wie im Zitat genannt, ein Allzweckwerkzeug, sondern ein für Ökonomen nützlicher Formalismus, um all die strategisch irrelevanten Details von einem Kontext abzustreifen, die für Psychologen eine zentrale Rolle spielen. Was also für die eine Wissenschaft nützlich ist, muss es nicht auch für eine andere sein. Die Neuroökonomie, und auch andere kognitiven Wissenschaften arbeiten laut Gul und Pesendorfer mit physiologischen und psychologischen Daten, weil diese für sie nützlich sind. Für ihre Forschungsfragen ist es wichtig, in welcher Weise sich Hirnregionen funktionell spezialisieren oder miteinander interagieren. Für die Arbeit von Ökonomen ist das ihrer Meinung nach nicht wichtig. Informationen über die Stärke von Insula-Aktivitäten sehen sie als irrelevant für die Beurteilung des Verhaltens von ökonomischen Agenten. Die Daten, mit denen Ökonomen arbeiten, sind nicht physiologisch, die Mittel nicht bildgebend. Die Daten und Mittel der Neuroökonomie schon. Nach Gul und Pesendorfer bearbeitet die Neuroökonomie Fragen mit Mitteln, die für die Beantwortung ökonomischer Fragen nicht interessant, nicht relevant sind. Die Neuroökonomie kann den Ökonomen keine interessanten, verwertbaren Erkenntnisse liefern, so ihre Meinung. Ein von den Neuroökonomien Camerer, Loewenstein und Prelec (Camerer et al. 2004, 573) formuliertes Ziel, die direkte Messbarkeit von Präferenzparametern durch neurowissenschaftliche

Mittel zu vereinfachen („asking the brain, not the person“), sehen sie als unrealistisch. Sie sind fest überzeugt, dass die Neuroökonomien kein Beispiel für die Beobachtung von Entscheidungsparametern durch die Anwendung bildgebender Verfahren vorweisen können und auch keine Ideen haben, wie das bewerkstelligt werden könnte. Es gibt ihrer Meinung nach keine Kriterien, um ein Gehirn beispielsweise bei einem Abschlagsfaktor von $\delta = 0,97$ von einem Gehirn bei $\delta = 0,7$ unterscheiden zu können.

Der Tenor bei Gul und Pesendorfer ist: Die Fragestellungen und Herangehensweisen der Neuroökonomie sind interessant und nützlich für die Neuroökonomie, aber für die Ökonomie haben sie keinen Nutzen, sie helfen Ökonomen bei ihren Fragestellungen nicht weiter. Schlimmer noch, selbst wenn Ökonomen Interesse an Gehirndaten zu ökonomischen Problemen hätten, könnten ihnen die Neuroökonomien keine brauchbaren Informationen liefern, die Neuroökonomie ist schlicht irrelevant für die Ökonomie. Das hat auch mit dem anfangs genannten zweiten Aspekt zu tun, dass Ökonomen keine neurowissenschaftlichen Probleme behandeln. Das ist ein zentrales Argument für Gul und Pesendorfer, das sich zu betrachten lohnt. Im Grunde lautet es zusammengefasst: Das ökonomische Standardmodell kann nicht mit physiologischen Gehirndaten widerlegt werden, denn die ökonomische Definition von Rationalität macht keinerlei Aussagen über physiologische Hirnprozesse. Oder, wie Gul und Pesendorfer (Gul et al. 2008, 21) es formulieren:

„Rationality in economics is not tied to physiological causes of behavior, and therefore, the physiological mechanisms cannot shed light on whether a choice is rational or not in the sense economists use the term. Brain mechanisms by themselves cannot offer evidence against transitivity of preferences or any other choice-theoretic assumption. Therefore, evidence that utility maximization is not a good model of the brain cannot refute economic models.“

Da *Expected Utility Theory* ausschließlich Vorhersagen über Entscheidungsverhalten macht, kann ihre Validität nach Gul und Pesendorfer auch ausschließlich anhand von Daten über Entscheidungen beurteilt werden. Physiologische Daten aus der Neuroökonomie sehen sie dafür als nutzlos. Daher können Neuroökonomien mithilfe ihrer neurowissenschaftlichen Studien auch nicht gegen ökonomische Modelle argumentieren. Anlass für dieses Argument sind Kritiken, die Gul und Pesendorfer bei Neuroökonomien finden, die sich darum drehen, dass das ökonomische Menschenbild unrealistisch ist, da sich die Prozesse, die es ökonomischen Agenten zuschreibt, nicht im Gehirn echter Menschen wiederfinden lassen. Sie beziehen sich auf zwei Textstellen von Rabin und Camerer. Rabin (Rabin 1998, 24) schreibt:

„Economists have traditionally assumed that, when faced with uncertainty, people correctly form their subjective probabilistic assessments according to the laws of probability. But researchers have documented many systematic departures from rationality in judgement under uncertainty.“

Während Rabins Textpassage keinen Bezug zu neurowissenschaftlichen Erkenntnissen nimmt (dafür ist sie auch bereits ein wenig zu alt), findet sich bei Camerer (Camerer 2005c, 1) die Idee zur Überprüfung ökonomischer Ideen durch die Neuroökonomie:

„For example, when economists think about gambling they assume that people combine the chance of winning (probability) with an expectation of how they will value winning and losing ('utilities'). If this theory is correct, neuroeconomics will find two processes in the brain – one for guessing how likely one is to win and lose, and another for evaluating the hedonic pleasure and pain of winning and losing and another brain region which combines probability and hedonic sensations. More likely, neuroeconomics will show that the desire or aversion to gamble is more complicated than that simple model.“

Gul und Pesendorfer sehen an dieser Textstelle wieder den Fehler der Neuroökonomien, die Herangehensweisen, Ziele und Mittel von Ökonomie und Neuroökonomie gleichzusetzen, um dann herauszufinden, welche Wissenschaft die korrekte Antwort liefert. Sie betonen, dass die Ökonomie aber nunmal keine Wissenschaft über das menschliche Gehirn ist (Gul et al. 2008, 4):

„The standard economic model of choice is treated as a model of the brain and found to be inadequate. Either economics is treated as amateur brain science and rejected as such, or brain evidence is treated as economic evidence to reject economic models.“

Die Ökonomie nach Gul und Pesendorfer macht keine Aussagen über das Gehirn, wie sie immer wieder betonen, und ist auch nicht dafür gedacht, Gehirnprozesse zu beschreiben oder zu erklären. Dass sich also die im ökonomischen Modell beschriebenen Prozesse nicht unbedingt 1:1 im menschlichen Gehirn wiederfinden lassen, finden Gul und Pesendorfer weder verwunderlich noch verwerflich. Rabin solle sich also nicht wundern, wenn das ökonomische Standardmodell kein gutes Modell des menschlichen Geistes ist, schließlich nehmen Ökonomen gar nicht an, dass Konsumenten in Wahrscheinlichkeiten denken oder den Satz von Bayes so gut kennen wie Wirtschaftsstudenten. Ökonomische Modelle sind ihrer Meinung nach durch ökonomische Variablen, Preise, Quantitäten und

anderem mit der Realität verbunden, nicht über ihre Modellage des Entscheidungsprozesses von Individuen.¹¹ Die Validität der ökonomischen Modelle wird laut den Autoren ohnehin ausschließlich mit Daten über Entscheidungen belegt. Die behavioralen Prozesse der menschlichen Entscheidungsfindung auf Regionen und Prozesse im Gehirn zu kartieren, ist in ihren Augen die Aufgabe von Neurowissenschaftlern und nicht von Ökonomen. Nach Meinung von Gul und Pesendorfer muss sich die Ökonomie also keine Gedanken darüber machen, ob ihr Modell eine gute Abbildung der tatsächlich im Gehirn vorgehenden Prozesse ist. Selbst wenn Neuroökonomien diese Prozesse in der Zukunft detailliert darstellen könnten, brauchen sich Ökonomen nicht damit zu beschäftigen, da die Ökonomie sich nicht damit beschäftigt, was im Gehirn von Individuen vorgeht. Eines zwar räumen sie Neuroökonomien für die Zukunft ein: Sie sehen die Möglichkeit, dass Erkenntnisse aus der Neuroökonomie Ökonomen inspirieren könnten, andere Modelle zu erstellen, aber ökonomische Modelle verwerfen, können sie nicht.

Gul und Pesendorfer sprechen ein paar spannende Aspekte an und sind sehr klar in ihren Aussagen. Dabei tritt hervor, welche Vorstellung sie von der Neuroökonomie und, noch interessanter, von der Ökonomie haben. Beispielsweise sind sie davon überzeugt, dass physiologische Daten für die Ökonomie irrelevant, nutzlos sind. Das liegt ihrer Meinung nicht nur an einem geringen Erkenntnisstand in der Neuroökonomie, sondern ist ein grundsätzlicher Zustand, der sich auch in Zukunft nicht ändern wird. Er resultiert für Gul und Pesendorfer aus den unterschiedlichen Zielen, die Ökonomie und Neuroökonomie mithilfe unterschiedlicher Mittel, anhand unterschiedlicher Fragestellungen verfolgen. Ökonomen können mit physiologischen Daten, salopp gesagt, in ihren Augen einfach nichts anfangen. Diese Einstellung teilen nicht alle Ökonomen. Auch solche, die keine ausgesprochenen Anhänger der Neuroökonomie sind, warnen davor, mögliches Potenzial so einfach vom Tisch zu fegen. Harrison und Ross (Harrison et al. 2010, 194) beispielsweise sehen die Neuroökonomie zwar durchaus kritisch, finden Guls und Pesendorfers Haltung jedoch zu drastisch:

„We reject the view that neural data are irrelevant to economics as willfully and obstructively isolationist. But we also reject the free-disposability view that *any* data are useful data until proven otherwise, implying that we should just collect the data anyway and decide later if they were useful. That is a poor model for advancement of study in any field. We welcome NE as a potential contributor to formal modeling of

¹¹ Das ist eine interessante Neuauflage des alten Friedman-Arguments, dass das ökonomische Modell nicht realistisch sein muss, da es als *As-if*-Modell konzipiert ist. (Vgl. oben, Abschn. 2.1)

the processes by which agents make economic decisions, though we emphasize that this project is in its infancy [...].“

Sie sehen durchaus nicht jede Datenquelle als nützlich, doch neuronale Daten von vornherein als nutzlos anzusehen, ist für Harrison und Ross eine hinderlich isolationistische Haltung. An einer anderen Stelle verwendet Harrison (Harrison 2008a, 339) in diesem Zusammenhang auch den Ausdruck „needlessly isolationist“. Tatsächlich bedeutet die Annahme, dass Neuroökonomie und Ökonomie unterschiedliche Fragestellungen, Mittel und Ziele haben, nicht notwendigerweise auch, dass die Ökonomie grundsätzlich niemals aus Erkenntnissen aus der Neuroökonomie wird profitieren können. Sicherlich gibt es eine Grenze zwischen Ökonomie und Kognitionswissenschaften und das nicht ohne Grund. Von vornherein aber Erkenntnisse aus dem jeweils anderen Gebiet als grundsätzlich und unveränderbar irrelevant zu behandeln, ist, da stimme ich mit Harrison überein, unnötig abweisend. Zwar räumen Gul und Pesendorfer in einem Halbsatz (Gul et al. 2008, 22; s. o. S. 159) ein, Erkenntnisse neuroökonomischer Art könnten Ökonomen dazu inspirieren, andere Modelle zu schreiben. Dieses Zugeständnis geht allerdings nicht so weit wie das theoretische Potenzial, das beispielsweise Harrison und Ross der Neuroökonomie zuschreiben. Es geht nicht einmal unbedingt von einer direkten Beeinflussung, einer direkten Nutzung gegenseitiger Erkenntnisse aus. Das ist zu schwach, um die Aussage von der Irrelevanz neurologischer Daten zu relativieren.

Diese Einstellung, keinen Einfluss aus den Kognitionswissenschaften zuzulassen, stimmt mit dem sich allgemein abzeichnenden Bild Guls und Pesendorfers von der Ökonomie überein. Sie lassen nicht nur keine neurologischen Daten als Erkenntnisquelle zu, sie sind auch der Meinung, dass es die Ökonomie im Grunde nicht zu interessieren braucht, warum handelnde Individuen die Präferenzen haben, die sie zeigen. Diese Ignoranz der Hintergründe zeichnet ein sehr konservativ neoklassisches Bild der Ökonomie, mit dem sich auch beispielsweise Friedman vor sechzig Jahren hätte identifizieren können.¹² Sie scheinen die gesamte Ökonomie als aus Standardökonomie, wie sie es nennen, bestehend zu sehen. Diese Sicht umfasst aber bereits seit Jahrzehnten nicht mehr die gesamte Wissenschaft Ökonomie, falls sie das jemals getan hat. Tatsächlich interessieren sich beispielsweise in der Verhaltensökonomie zahlreiche Ökonomen für die Frage, warum Individuen die Präferenzen zeigen, die sie zeigen, beziehungsweise welche Motive sie für ihre Entscheidungen haben. Mit dem Ausschluss solcher Fragen ignorieren Gul und Pesendorfer einen ganzen Teilbereich der Ökonomie,

¹² An dieser Stelle sei an Friedmans Argumente und Zitate in Abschnitt 2.1 erinnert.

der nicht mit ihrer Auffassung von den Zielen, Mitteln und Daten zeitgenössischer Forschung übereinstimmt. Da Guls und Pesendorfers Argumentation, wie wir gesehen haben, aber auf diesen Annahmen, diesem Bild der Ökonomie aufbaut, können sie auch nicht für diejenigen Ökonomen sprechen, die in der Tat an solchen Fragen forschen. Stattdessen können sie nur diejenigen Ökonomen vertreten, die wie sie selbst einer konservativ neoklassischen Standardökonomie angehören, die an verhaltensökonomischen Zielen, Mitteln und Daten nicht nur nicht interessiert sind, sondern sie sogar ablehnen, da sie sie offenbar nicht als Teil der Ökonomie ansehen. Als ökonomischer Leser, der vermutlich die hauptsächliche Zielgruppe von Guls und Pesendorfers Artikel darstellt, kann man ihren Argumenten also zustimmen, muss es aber nicht. Man kann ihre Argumente mit Leichtigkeit nicht teilen, wenn man nicht mit ihren Prämissen, in diesem Fall ihrem Bild von der modernen Ökonomie, übereinstimmt. Welche gravierenden Folgen das für die neuroökonomische Zusammenarbeit hat, werden wir unten in Abschnitt 6.2.3 sehen.

Zunächst aber wenden wir uns noch einem anderen Punkt zu, der bei der Lektüre des Artikels auffällt und weitere Fragen aufwirft bezüglich der gewählten Argumentationsstrategie der beiden Autoren: Ein Grund, den Gul und Pesendorfer dafür nennen, dass physiologische Daten keine Relevanz für die Ökonomie haben, ist, dass die Standardökonomie andere Ziele mit anderen Mitteln verfolgt als die Neuroökonomie, und die Ökonomie keine physiologischen Annahmen macht. Das ökonomische Standardmodell macht keine Angaben darüber, was im Gehirn der handelnden Individuen vorgeht. Es wurde auch nicht auf physiologischen Überlegungen begründet und kommt ohne Informationen aus den Neurowissenschaften aus, was Gul und Pesendorfer auch als Grund dafür ansehen, dass neuronale Daten das ökonomische Modell nicht verwerfen können. Daher sagen sie:

„Rationality in economics is not tied to physiological causes of behavior, and therefore, the physiological mechanisms cannot shed light on whether a choice is rational or not in the sense economists use the term.“ (Gul et al. 2008, 21)

So weit, so bekannt. Interessanterweise sagen sie zu diesem Thema auch:

„Neuroscience evidence cannot refute economic models because the latter make no assumptions and draw no conclusions about the physiology of the brain. Conversely, brain science cannot revolutionize economics because the former has no vehicle for addressing the concerns of economics.“ (Gul et al. 2008, 4)

Die Kognitionswissenschaften haben kein „vehicle“, wie sie es nennen (Gul et al. 2008, 22), um die Probleme der Ökonomie anzusprechen. Diese Formulierung führen Gul und Pesendorfer nicht weiter aus, dabei wird es an dieser Stelle erst richtig interessant. Was meinen sie genau mit „vehicle“? Hier hätten sie grundsätzliche, wissenschaftstheoretische Probleme bei der Zusammenarbeit zweier oder mehrerer Wissenschaften ansprechen können, sogar für reduktionistische und antireduktionistische Bedenken wäre hier ein Ansatzpunkt gewesen. Sie sprechen in ihrem Aufsatz durchaus an, dass sie bei den Neuroökonomern den Wunsch nach einer „unification“ (Gul et al. 2008, 32) der Neurowissenschaften mit der Ökonomie sehen (Gul et al. 2008, 32–34). Daher sprechen sie wiederholt und vehement davon, dass die beiden Wissenschaften unterschiedliche Fragestellungen, Ziele und Mittel haben und unterschiedliche Daten verarbeiten müssen und es daher nie zu einer *Unification* kommen kann. Diese Argumentation bleibt aber nur an der Spitze des Eisberges, der erscheint, wenn man die Frage nach dem *Vehicle* stellt, mit dem die beiden Wissenschaften verbunden werden können. Gul und Pesendorfer argumentieren zwar in die Richtung ‚keine Reduktion der Ökonomie auf die Neurowissenschaften‘, aber sie stoßen diese Türe nicht auf. Sie sprechen nie von Reduktion und verwenden auch kein Vokabular aus diesem Bereich. Dabei wäre eine solche Diskussion nicht zu weit hergeholt, denn in seiner 2011 veröffentlichten Monografie *Foundations of Neuroeconomic Analysis* spricht Glimcher explizit über eine mögliche Reduktion der Ökonomie auf die Neurowissenschaften, am besten mithilfe von Brückengesetzen. Dieses Werk erschien zwar drei Jahre nach dem Artikel von Gul und Pesendorfer, die Idee wird jedoch nicht neu gewesen sein, vor allem, da Gul und Pesendorfer sie ja bereits Jahre zuvor in neuroökonomischen Aufsätzen entdeckt haben wollen. Es sollte eigentlich eine günstige Gelegenheit für alle Gegner der Neuroökonomie sein, wenn prominente Neuroökonomern selbst von Reduktionsmöglichkeiten sprechen und sogar schon nach Brückengesetzen suchen. Denn sollten Neuroökonomern im Reduktionismus mehr sehen als eine rein metaphysische Position, müssen sie sich auch mit den seit Jahrzehnten diskutierten Problemen des Reduktionismus auseinandersetzen. Um die neuroökonomischen Ideen von einer *Unification* und damit auch die Verwerfung ökonomischer Modelle durch neuroökonomische Erkenntnisse wirksam grundlegend zu kritisieren, hätten Gul und Pesendorfer also nicht ihre recht schwammige Argumentation von den unterschiedlichen Zielsetzungen vorbringen müssen. Altbekannte Argumente gegen den Reduktionismus hätten eine größere Wirkung erzielt.

Also warum haben Gul und Pesendorfer diese Möglichkeit nicht genutzt? Entweder sie sind über die reduktionistische Idee der *Unification of the Sciences* nicht

sehr gut informiert oder sie wollten dessen Argumente absichtlich nicht verwenden. Bei letzterer Möglichkeit stellt sich natürlich die Frage, warum sie das nicht wollten. Da sie beide Ökonomen sind und ihr Publikum, dem Tonus des Aufsatzes nach zu schließen, vor allem ebenfalls aus Ökonomen besteht, wollten sie die Diskussion eventuell nicht auf eine den Ökonomen fremde wissenschaftstheoretische Ebene bringen. Diese Hypothese ist allerdings reine Spekulation. Der Grund für Gul und Pesendorfers Verzicht auf wissenschaftstheoretisches Vokabular ist aber nicht unwichtig, vor allem im Hinblick auf die folgenden Überlegungen.

6.2.3 No Brain in Economics

In diesem Abschnitt geht es nicht um ein bestimmtes Argument, das Gul und Pesendorfer gegen die Neuroökonomie vorbringen, sondern genauer als bisher um ihre allgemeine Sicht ihrer Wissenschaft. Das Verständnis davon, wie die beiden Ökonomen die grundlegende Konzeption ihrer eigenen Wissenschaft sehen, hilft wiederum dabei, die Probleme, die Meinungsverschiedenheiten zwischen Ökonomen und Neuroökonomien zu verstehen. Wir werden sehen, dass diese grundsätzlicher Natur sind und nur unter ganz bestimmten Umständen ausgeräumt werden können. Ökonomen wie Gul und Pesendorfer haben eine von den Neuroökonomien komplett verschiedene Ansicht davon, wie Ökonomie arbeitet und welchen Fragen sie nachgeht. Wie wir sehen werden, scheint es, als würden beide unter verschiedenen Paradigmen arbeiten und daher nicht zusammenfinden.

Gul und Pesendorfer vertreten, wie wir oben gesehen haben, eine Sicht der Ökonomie, die die Verwendung neurowissenschaftlicher Daten ablehnt und keinen Nutzen für Erkenntnisse aus den Kognitionswissenschaften für die Ökonomie ausmachen kann. Damit stehen sie in der Tradition von Ökonomen wie Friedman, der, wie wir in Abschnitt 2.2 gesehen haben, die Unabhängigkeit ökonomischer Modelle von unbeobachtbaren psychologischen Vorgängen in den Köpfen von Individuen als ihre große Stärke sieht und an dieser Art Modell festhalten will. Tatsächlich ist die Simplizität der *Revealed-Preference*-Axiome einer der Gründe, warum sie so erfolgreich und vielfältig anwendbar sind. Physiologische Erkenntnisse in ökonomische Modelle einzuarbeiten, könnte sie komplexer und damit komplizierter machen, noch dazu können weitere Annahmen die Falsifizierbarkeit erhöhen. In den Augen von Standardökonomien (übernehmen wir diese Vokabel) wie Gul und Pesendorfer wäre es also reichlich unnützlich, das Standardmodell komplexer zu machen und sich nicht zuletzt der Problematik auszusetzen, dass die nicht direkt beobachtbaren Hirnprozesse von Individuen das Modell anfälliger für Fehler werden lassen. Wie aus ihrem Aufsatz klar wird, ist ihr Vertrauen in die

Stärke der kognitiven Neurowissenschaften, zumindest derzeit, nicht sehr groß, sie sehen bei ihnen noch keine brauchbaren Erkenntnisse (Gul et al. 34; s. o. S. 157). In den Augen der StandardökonomInnen braucht die Ökonomie die Neuroökonomie und ihre Erkenntnisse nicht. Denn die Ökonomie kommt ihrer Meinung nach sehr gut (oder besser) mit der Annahme des menschlichen Gehirns als Black Box zurecht. Für ihre Arbeit sind die Präferenzen von Individuen wichtig, und um diese zu ermitteln, gibt es für StandardökonomInnen keine bessere Methode als das *Revealed-Preference*-Modell. Denn mit einer Black Box als Gehirn gibt es nicht mehr Informationen über Individuen einzusammeln als eine Liste ihrer Präferenzen. Mit dem *Revealed-Preference*-Modell klappt das nach Meinung der StandardökonomInnen ganz hervorragend. Die Black Box zu öffnen und schwer nachvollziehbare physiologische Daten mit in das Modell aufzunehmen, wäre daher nicht nur unnötig, es würde das Modell und damit die Leistungsfähigkeit der Ökonomie zum Schlechteren verändern. Diese Einstellung ist eindeutig behavioristisch. Methodologischer Behaviorismus zeichnet sich dadurch aus, dass er mentale Zustände nicht als Erkenntnisquelle für die Erforschung des Verhaltens von Individuen zulässt, da sie nicht von außen beobachtbare Entitäten sind. Das entspricht Guls und Pesendorfers Ansicht davon, wie Ökonomie als Wissenschaft funktioniert und weiterhin funktionieren soll. Erkenntnisse aus den Kognitionswissenschaften helfen der Ökonomie in ihren Augen nicht dabei, das Verhalten von Individuen zu modellieren. Kognitionswissenschaften dagegen, zu denen neben den kognitiven Neurowissenschaften auch die kognitive Psychologie gehört, sind in den vergangenen Jahrzehnten zu ihrer eigenen Erleichterung dem methodologischen Behaviorismus entwachsen und sehen von außen unbeobachtbare mentale Prozesse naturgemäß als wichtige Erkenntnisquelle zum Verständnis menschlichen Verhaltens. In diesen neu erschlossenen Möglichkeiten sehen sie ein großes Potenzial auch für die Ökonomie, das die Ökonomen in ihrer Meinung nach nicht ungenutzt lassen sollten. Es gibt also zwischen NeuroökonomInnen und StandardökonomInnen grundlegende konzeptionelle Verschiedenheiten, was die Anschauung der Ökonomie als Wissenschaft angeht. Die einen sehen neurologische Forschung als bedeutenden Meilenstein auch für andere Wissenschaften, während die anderen neurologische Erkenntnisse als Informationsquelle per se ausschließen. Diese Anschauungsunterschiede sind nicht miteinander vereinbar. Es würde in diesem Fall nicht helfen, gemeinsame Konventionen und Forschungsstandards festzulegen und die Kommunikation zu erhöhen. Solange die StandardökonomInnen an ihrer behavioristischen Sichtweise festhalten, werden sie eine Kooperation mit den Neurowissenschaften weder gutheißen wollen, noch können, denn für sie, in ihrer wissenschaftlichen Umgebung, sind die neurologischen Erkenntnisse nutzlos, ja sogar schädlich. Sie erkennen die Validität

neuroökonomischer Daten und Erkenntnisse nicht an. Das bedeutet andererseits, dass diejenigen Ökonomen, die keine behavioristische Sicht auf die Ökonomie haben und die Validität kognitionswissenschaftlicher Erkenntnisse nicht bestreiten, durchaus grundsätzlich an einer ernsthaften Kooperation mit den Neurowissenschaften interessiert sein könnten, beispielsweise mit einer ähnlichen Einstellung wie Harrison und Ross. Die schlechte Nachricht dabei ist allerdings, dass Neuroökonomien bei denjenigen Ökonomen, die den Behaviorismus als einzig richtige Theorie für die Ökonomie sehen, nie auf Zustimmung werden stoßen können. Ihre Ansichten, ihre Art, Wissenschaft zu betreiben, sind inkompatibel und werden es bleiben. Dieses Problem kann nicht behoben werden, solange nicht eine der beiden Seiten ihren Standpunkt ändert. Das ist auch der Grund warum Ökonomen wie Gul und Pesendorfer die Neuroökonomie für eine nicht durchführbare, unverständliche Idee halten. Sie verstehen nicht, wie Neuroökonomien neurologische oder psychologische Daten für eine wissenschaftlich verwertbare Erkenntnisquelle halten können, ebensowenig wie Neuroökonomien verstehen, wieso manche Ökonomen diese vielversprechende neue Datenquelle nicht nutzen möchten. Diese Situation erinnert an Kuhn (Kuhn 1962) und seine Idee von der Inkommensurabilität wissenschaftlicher Theorien. Das ist eine kontroverse Idee, die aber dabei hilft, zu verstehen, was das Problem von Gul und Pesendorfer und aller Standardökonomien mit der Neuroökonomie ist. Da Inkommensurabilität auch ein Argument gegen Reduktionismus ist und ein wenig mehr Platz benötigt, als an dieser Stelle vorhanden ist, werden wir uns mit diesem Thema weiter unten in Abschnitt 6.3 (S. 180) eingehender befassen.

Die Erkenntnis, dass Gul und Pesendorfer eine behavioristische Auffassung von der Ökonomie vertreten und diese per Definition schon nicht mit Kognitionswissenschaften vereinbar ist, ist recht simpel. Sie ist weder schwer zu erkennen, noch zu verstehen. Daher kommt auch an dieser Stelle die Frage auf, warum Gul und Pesendorfer ihr Problem mit der Neuroökonomie nicht auch so klar formulieren. Damit könnten sie nicht nur ihre Probleme klar darstellen, sondern auch allen anderen Standardökonomien begreiflich machen, dass für sie eine Kooperation mit der Neuroökonomie schon konzeptionell gar nicht möglich ist. Diese einfache Darstellung des Problems haben sie ungenutzt gelassen. Warum? Haben sie diese Hypothese durchdacht und verworfen? Sehen sie sich nicht in behavioristischer Tradition? Das wird sich aber kaum bestreiten lassen. Oder ist ihnen die behavioristische Einstellung nicht bewusst? Vokabular aus diesem Bereich taucht jedenfalls in ihrem Aufsatz nicht auf (und soweit mir bekannt ist, auch in keinem späteren). Auf dieser Grundlage kann kein Urteil darüber gefällt werden, ob sie die wissenschaftstheoretischen Aspekte ihrer eigenen Arbeit nicht erkannt

haben oder ob sie durchaus darum wissen, sich aber bewusst dagegen entschieden haben, auf dieser Grundlage zu argumentieren. Stattdessen haben sie sich dafür entschieden, auf 36 Seiten zu kritisieren, dass Neuroökonomien das ökonomische Wohlfahrtskriterium nicht verstanden haben und nicht einsehen wollen, dass sie andere Forschungsfragen verfolgen als Ökonomen. Das war nicht unbedingt die bessere Wahl, denn diese Argumente sind schwammiger und zeigen lediglich äußerlich sichtbare Symptome des darunterliegenden prinzipiellen Problems auf. Sie haben die Chance nicht genutzt, die komplexe Situation einfach zusammenzufassen. Dabei wäre damit die Grundlage der systematischen Kritiken geklärt, die davon handeln, dass die Neuroökonomie der Ökonomie nicht weiterhilft. Guls und Pesendorfers Argumente ließen sich auf diese Erkenntnis von der Unvereinbarkeit reduzieren.

Danach bliebe immer noch die große Frage, welche Seite die Ökonomie als Ganzes einnehmen sollte. Sollte sie an behavioristischen Methoden festhalten oder sich kognitionswissenschaftliche und eventuell weitere Erkenntnisse zunutze machen? Das ist allerdings keine spezifisch neuroökonomische Frage, sondern eine intern ökonomische, die die Ökonomie nun bereits seit Jahrzehnten beschäftigt, wie wir in den Kapiteln 2 und 3 gesehen haben. Über diese Frage werden sich Standardökonomien und Neuroökonomien vermutlich noch viele weitere Jahre streiten, allerdings sind Neuroökonomien nicht allein aufgrund ihrer spezifischen Eigenschaft als Neuroökonomien an diesem Streit beteiligt, sondern weil sie wie viele andere Ökonomen keine behavioristischen Ansichten teilen. Man muss nicht Neuroökonom sein, um sich gegen eine behavioristische Ausrichtung der Ökonomie, wie Gul und Pesendorfer sie vertreten, auszusprechen. Psychologische Forschung beispielsweise war in einem Großteil des 20. Jahrhunderts durch den Behaviorismus geprägt. Das mag zu einem Fortschritt auf dem Gebiet der Laborexperimente geführt haben, doch schränkte es die Forschungsmöglichkeiten von Psychologen massiv ein. Alles, was nicht direkt beobachtbar ist, war nicht als Datenquelle zugelassen. Gewisse Fragen, wie über Bewusstsein, konnten somit gar nicht bearbeitet werden. Nicht wenige Psychologen waren nicht traurig, als der Behaviorismus in der zweiten Hälfte des Jahrhunderts an Bedeutung verlor und sich ganz neue Felder wie die kognitive Psychologie etablieren konnten. Neue Erkenntnisquellen standen nun offen und haben neue Sichtweisen auf menschliches Handeln hervorgebracht, die unter behavioristischen Regeln nicht möglich gewesen und weiter unbekannt geblieben wären, hätte die wissenschaftliche psychologische Gemeinschaft weiterhin am Behaviorismus festgehalten. Ähnlich könnte es der Ökonomie ergehen, wenn sie am Behaviorismus festhält. Moderne Verhaltensökonomien wie Selten versuchen seit Jahrzehnten, diese Grundeinstellung unter Ökonomen aufzubrechen und neue Erkenntnisquellen zu etablieren

(vgl. Güth et al. 1997, 4, 6). Damit haben sie heute mehr Erfolg als noch vor dreißig Jahren, aber noch immer gibt es Argumentationen wie die betrachtete von Gul und Pesendorfer, die sich tatsächlich als unnötig isolierend bezeichnen lässt. Es ist sicher ein einfaches und erfolgreiches Rezept, sich nur auf Datenquellen zu verlassen, die beobachtbare Daten liefern, statt Spekulationen auf Konstruktbasis anzustellen. Aus diesen Regeln Schranken zu machen, über die die ökonomische Forschung nicht hinaustreten darf, versperrt allerdings schnell den Weg zu weiteren, wertvollen Erkenntnissen.

Das heißt allerdings nicht, dass Guls und Pesendorfers Argumente gegen eine Zusammenarbeit von Ökonomen und kognitiven Neurowissenschaftlern verpuffen. Sie sprechen durchaus problematische Punkte an, die auch nicht an Brisanz verlieren, wenn man die behavioristische Einstellung der Autoren nicht teilt. Diese Probleme werden wir im folgenden Abschnitt 6.3 näher betrachten.

6.3 Wissenschaftstheoretische Kritiken an der Neuroökonomie

Wir haben uns mit den Kritiken von Gul und Pesendorfer beschäftigt, die sich vor allem mit grundlegenden systematischen Problemen zwischen Neuroökonomien und Standardökonomien auseinandersetzen. Solche Kritiken sind grundsätzlicher Natur und daher spannender als viele Kritiken, die sich vermutlich im Laufe der Zeit von allein entschärfen, wie beispielsweise die Kritik an zu schlechten Messtechniken aus Abschnitt 6.1.2. Gul und Pesendorfer sind zwei von wenigen Autoren, die solche systematischen Probleme besprechen. Mit ihren Aufsätzen zum Thema hauen sie, bildlich gesprochen, mit der Faust auf den Tisch. Das erzeugt großen Krach und viel Aufmerksamkeit, ist aber, auch das haben wir gesehen, eher schlecht gemacht. Sie lassen die ihren Argumenten zugrundeliegenden wissenschaftstheoretischen Konzeptionen außen vor und vertun damit die Chance, ihre Ideen und Bedenken auf argumentativ sichere Beine zu stellen. Denn die Probleme, von denen sie sprechen, wurden bereits vor Jahrzehnten, beziehungsweise auch heute noch, in wissenschaftstheoretischen Diskussionen behandelt. Statt die Probleme bei der Wurzel zu packen und ihre Ursachen zu klären, beschreiben sie lediglich deren Symptome. Das spricht gegen die Vorgehensweise von Gul und Pesendorfer; aber spricht es auch für die Neuroökonomie? Bedeutet die Einsicht, dass Gul und Pesendorfer ihre Kritiken nicht gut fundiert haben, auch, dass die Neuroökonomie jeder Kritik enthoben ist? Sind ihre Kritikpunkte damit ausgeräumt, bleibt nichts von ihnen übrig? Doch, tatsächlich bleiben Kritikpunkte bestehen, auch wenn Gul und Pesendorfer auf

Chancen verzichtet haben, sie zu formulieren. Diese werden wir im Folgenden genauer betrachten. Vor allem die Frage nach der Einstellung der Neuroökonomien zum Reduktionismus ist einer genaueren Betrachtung wert. Wollen verschiedene Wissenschaften wie Ökonomie und Neurowissenschaften miteinander kooperieren und von gegenseitigen Erkenntnissen profitieren, sollten sie auf irgendeine Weise miteinander verknüpfbar, kompatibel sein. Ökonomische Modelle machen, wie Gul und Pesendorfer argumentieren, keine Aussagen über das menschliche Gehirn; sie kommen ohne solche Aussagen aus. Dennoch sollen sie in einer neuroökonomischen Zusammenarbeit von Erkenntnissen profitieren, die genau solche Aussagen trifft. Wie aber sollen ökonomische Modelle und Theorien, denen Gehirndaten fremd sind, diese Erkenntnisse überhaupt aufnehmen oder verarbeiten? Solche Probleme wurden von Anhängern des Reduktionismus so gelöst, dass beide Wissenschaften auf dieselbe Basis gebracht werden müssen, damit sie kompatibel sind oder eine Theorie aus einer anderen, basaleren abgeleitet werden kann. Zu diesem Thema gab es im vergangenen Jahrhundert viele Modelle und ebenso viel Kritik von betroffenen Wissenschaftlern und Wissenschaftstheoretikern. Diese Diskussionen sind im Umfeld der Neuroökonomie heute wieder relevant, denn vieles, was Gul und Pesendorfer ansprechen, wurde von Philosophen, Psychologen und Physikern damals bearbeitet. Daher werden wir uns im Folgenden mit ihren Kritiken und Diskussionen befassen, beginnend bei reduktionistischen Methoden und Problemen, denen sich heutige Neuroökonomien entgehen sehen. Dazu betrachten wir zunächst am Beispiel von Glimcher, wie ein bekannter Neuroökonom sich die wissenschaftstheoretisch nicht unproblematische Zusammenarbeit von Neurowissenschaftlern und Ökonomen vorstellt. Daran werden wir sehen, welchen Problemen sich seine Konzeptionen stellen müssen und dass diese sehr ähnlich den Problemen sind, die bereits von Ernest Nagel, Kenneth Schaffner, Paul Feyerabend und Fodor seit den 1960er Jahren verhandelt werden.

Wie oben bereits angesprochen, stellt Glimcher (Glimcher 2011) konkrete Überlegungen zur Verknüpfung von Ökonomie und Neurowissenschaften an.¹³ Er stellt zunächst die Ideen von Nagel (Nagel 1961) und Vertretern des logischen Positivismus vor und erläutert ihre Vorstellungen von vollständiger Reduktion auf die Physik. Er macht klar, dass er sich in ontologischer Opposition zu einer vollständigen Reduktion der Ökonomie auf die Neurowissenschaften sieht (Glimcher 2011, 25):

¹³ Siehe oben S. 162.

„[...] I would describe my own stance as an *ontological* opposition to complete reductionism. [...] I do not believe that all of economics can be reduced to neuroscience. This is actually a very important point, and it is why my own scholarly career has been devoted to the study of economics, psychology, and neuroscience.“

Damit distanziert sich Glimcher von der heute weithin als gescheitert erachteten reduktionistischen Idee von der *Unity of the Sciences*. Warum aber erst eine gescheiterte Idee vorstellen und sich dann von ihr distanzieren? Wenn er mit ihr ohnehin nicht übereinstimmt und sie auch heute nicht für anwendbar auf Ökonomie und Neurowissenschaften hält, warum sich überhaupt die Zeit nehmen, sie zu erläutern? Es drängt sich der Gedanke auf, dass Glimcher diesen Teil des Kapitels nicht nur als historischen Unterbau für seine folgenden Ausführungen benötigt, sondern auch den rhetorischen Kniff anwendet, eine heute abgelegte Idee als abwegig und gescheitert aufzubauen. Vor diesem Hintergrund könnte den Lesern seine Alternative dazu, die er darauffolgend vorstellt, als weitaus weniger abwegig erscheinen. Zusätzlich verortet er explizit Guls und Pesendorfers Argumente als Argumente gegen eine vollständige Reduktion (Glimcher 2011, 21). Distanziert er sich von dieser, distanziert er sich auch von Guls und Pesendorfers Argumenten. Doch wie wir sehen werden, entkommt er den Problemen des Reduktionismus nicht komplett, denn die von ihm vorgestellte Alternative übernimmt doch mehr von Nagels Ideen, als er seinen Lesern verrät. Betrachten wir sein weiteres Vorgehen:

Glimcher hält fest, dass sich jeder reduktionistische Versuch, die Sozial- und die Naturwissenschaften zu verbinden, zwei entscheidenden Fragen stellen muss (Glimcher 2011, 30):

„Are all concepts at the level of economics homomorphic to concepts at the neurobiological level?

Are all concepts at the level of economics emergent, or is it the case that some concepts are reducible (as is, or after modification) while others are emergent?“

Glimcher betont, dass wir, auf einem metaphysischen Level die Antworten auf diese Fragen nicht kennen. Seiner Meinung nach kann man nicht mit absoluter Gewissheit sagen, dass in ferner Zukunft ein Teil der Biologie auf die Chemie reduzierbar sein wird. Andererseits könnte sich laut Glimcher auch alles als reduzierbar herausstellen. Wie sollte also weiter vorzugehen sein, will man Ökonomie und Neurowissenschaften auf reduktionistische Weise verbinden? Glimcher beantwortet diese Frage mit einem Verweis auf Fodor (Fodor 1974), der laut Glimcher argumentiert hatte, dass sich die Psychologie vermutlich nicht auf die Biologie reduzieren lassen werde, da sie im Kern eine emergente Wissenschaft

ist.¹⁴ Seiner Meinung nach könnten psychologische Konzepte wie Aufmerksamkeit, Emotionen und Grammatik nicht auf neurobiologische Arten („kinds“, Glimcher 2011, 30) übertragen werden. Glimcher betont, dass sich diese Vorhersagen Fodors inzwischen als falsch herausgestellt haben. Das macht er daran fest, dass Psychologen heute weltweit mit Gehirncannern arbeiten. Aufmerksamkeit lässt sich nach Glimcher scheinbar leicht auf Hirnregionen kartieren, und Theorien aus Psychologie und Neurowissenschaften enthalten heute seiner Einschätzung nach viele gleiche logische Objekte und reduktive Verbindungen (Glimcher 2011, 31). Was das bedeutet, oder an welche Beispiele der dabei denkt, sagt er leider nicht. Er räumt zwar ein, dass die Kartierung von Emotionen auch heute nicht so einfach ist und sich auch Schlüsselkonzepte aus der Erforschung von Grammatik bisher einer Reduktion gänzlich entziehen. Aus diesem historischen Rückblick schließt Glimcher aber auf eine geeignete Vorgehensweise für die reduktive Verbindung von Ökonomie und Neurowissenschaften (Glimcher 2011, 31):

„This history, then, tells us something fairly clear that can guide us as we try to relate neurobiology, psychology, and economics. At an empirical level there almost certainly will be regularities that homomorphically map some economic kinds to neurobiological kinds. [...] It also seems very likely that there will be emergent concepts in economics that resist reduction. What I am saying here is that the reductive relationship between neurobiology and economics (largely via psychology) will be incomplete – but only a real extremist could argue today that no such mappings of any kind will ever be found to exist. So far, every scientific discipline that has ever faced a reductive challenge (whether it be chemistry, biology, psychology, or anything else) has demonstrated at least a partial set of linkages to the more reduced disciplines. These partial reductions have also been influential; they have reshaped the logical kinds of the higher-level discipline. There is no reason to believe that economics will be different.“

Glimchers Meinung nach ermöglicht die Neuroökonomie den Aufbau einer interdisziplinären Synthese aus Ökonomie und Neurowissenschaften, die er nicht nur für prinzipiell möglich, sondern auch für höchstwahrscheinlich fruchtbar für die beteiligten Wissenschaften hält. Um eine Synthese erreichen zu können, sagt Glimcher, benötigt er reduktive Verbindungen („reductive linkages“, Glimcher 2011, 33), die in der gegenwärtigen und zukünftigen Forschung gesucht werden müssen. Leider sagt er nicht, was Verbindungen zu reduktiven Verbindungen

¹⁴ Wir werden den Emergenz-Begriff ab S. 174 betrachten, wenn wir auf die Probleme des Reduktionismus zu sprechen kommen.

macht. Allerdings sieht er hier das Problem, dass es keine reduktiven Verbindungen in der Welt gibt, die neurobiologisch, psychologisch und ökonomisch beobachtbare Phänomene direkt und ohne Modifikation verbinden. Daher gilt für Glimcher (Glimcher 2011, 32):

„[...] neuroeconomics would not be simply the linkage of three existing bodies of theory: neuroeconomics would be a revision of three existing bodies of theory into a more powerful linkable form.“

Denn (Glimcher 2011, 33):

„Our existing theoretical frameworks in all three disciplines are imperfect. That means that searching for alterations to each of those frameworks that allow linkages allows us to improve each of those sets of theories. If we believe partial reduction is possible in an ontological sense, then that means we can use knowledge gained in one field to constrain the structure of knowledge in the other fields. It means that knowing how the brain works, for example, will help us to constrain economic theory.“

Glimcher versucht also, reduktive Verbindungen zwischen drei Disziplinen zu finden. Diese drei Disziplinen verwenden unterschiedliches Vokabular und sind daher nicht ohne Weiteres reduktiv verbindbar. Glimcher appelliert daher (Glimcher 2011, 35):

„If we are trying to build principled linkages between these fields that will have real theoretical traction, we must do so in a serious manner. The critical idea here is that our goal must be to modify the theories of each of the parent disciplines so that they align toward a common endpoint and share as many reducible logical kinds as possible.“

Zusammengefasst beschreibt Glimcher hier als Ziel eine partielle Theorienreduktion, die er, im Gegensatz zur vollständigen Reduktion, in einem ontologischen Sinne für möglich hält. Dabei befinden sich Ökonomie, Neurobiologie und Psychologie in einer heterogenen Konstellation, da sie sich, wie Glimcher selbst feststellt, auf unterschiedliches Vokabular stützen. Die reduktiven Verbindungen zwischen den drei Wissenschaften sind also nicht trivial vorausgesetzt, sondern benötigen Brückengesetze. Glimcher erläutert in obigem Zitat die Idee, die Theorien aus Ökonomie, Psychologie und Neurowissenschaften auf eine Weise zu modifizieren, dass sie auf einen gemeinsamen Endpunkt ausgerichtet sind und so viele *logical Kinds* wie möglich gemeinsam haben. Mit dieser Idee hat er sich offenbar an Schaffner orientiert, der 1967 (Schaffner 1967) die Methode vorbrachte, die an einer Reduktion beteiligten Theorien zu modifizieren, um

eine Ableitung der übergeordneten Theorie von der untergeordneten Theorie zu ermöglichen. Um diese Idee und damit auch Glimchers Idee besser verstehen zu können, betrachten wir Schaffners Konzept in groben Zügen ein wenig näher, bevor wir mit den Kritiken am Reduktionismus und damit auch an Glimchers Idee von der Neuroökonomie fortfahren. Dazu beginnen wir ein paar Jahre vor Schaffner mit einer Kritik von Feyerabend:

Mit seinen Überlegungen antwortete Schaffner auf den ein paar Jahre zuvor von Feyerabend (Feyerabend 1962) vorgebrachten Einwand, bei einer Reduktion würden nicht die Gesetze der zu reduzierenden Theorie abgeleitet, sondern etwas, das ähnlich aussieht, aber eine andere Bedeutung hat. Nach Feyerabend verändert der Reduktionsprozess die Bedeutung der theoretischen Begriffe und die Bedeutung wenigstens einiger der Beobachtungsbegriffe einer zu reduzierenden Theorie, was seiner Meinung nach gravierende Folgen hätte (Feyerabend 1962, 29):

„That is, not only will description of things and processes in the domain in which so far T' [zu reduzierende Theorie] had been applied be infiltrated, either with the formalism and the *terms* of T [reduzierende Theorie], or if the terms of T' are still in use, with the *meanings* of the terms of T, but the sentences expressing what is accessible to direct observation inside this domain will now mean something different.“
[Anmerkungen von mir, Hervorhebungen im Original]

Daher verknüpfen die Brückengesetze laut Feyerabend nicht wie intendiert die Begriffe von T mit denen von T', sondern lediglich mit modifizierten Begriffen von T', die ähnlich aussehen mögen, aber eine andere Bedeutung haben. Das wäre insofern problematisch, als dass bei einer dennoch durchgeführten Reduktion die zu reduzierende Theorie verfälscht würde. Die ursprünglichen Begriffe der zu reduzierenden Theorie verschwänden dann, da die aus der reduzierenden Theorie abgeleiteten Begriffe eine andere Bedeutung hätten.

Dieses Problem konnten die Anhänger der Reduktion nicht ignorieren. Schaffner stellt sein Konzept von Reduktion vor, in dem die zu reduzierende Theorie modifiziert wird, um reduzierbar zu werden. In groben Zügen funktioniert das folgendermaßen: T_2 sei eine sekundäre, zu reduzierende Theorie, und T_1 sei eine primäre Theorie, auf die T_2 reduziert werden soll. Nun muss von T_2 eine Theorie T_2^* abgeleitet werden. Das Verhältnis der beiden soll eine starke Analogie im Sinne einer starken Ähnlichkeit sein. Zudem soll T_2^* T_2 korrigieren, indem sie akkuratere, in Experimenten verifizierbare Vorhersagen trifft und offenlegt, warum T_2 nicht korrekt war. Alle Begriffe aus T_2^* müssen von Brückengesetzen

durch die Begriffe aus T_1 definiert werden. Auf diese Weise kann T_2^* deduktiv aus T_1 abgeleitet werden, wenn T_1 mithilfe einer Anzahl an Brückengesetzen verbunden wird, die Schaffner „reduction functions“ (Schaffner 1967, 142) nennt.

In diesem Konzept muss die zu reduzierende Theorie modifiziert werden, damit eine Reduktion überhaupt möglich ist. Der Punkt dabei ist, dass T_2^* so formuliert werden soll, dass sie nur Begriffe enthält, die auch in der reduzierenden Theorie T_1 plus den Brückengesetzen enthalten sind, wodurch die Reduktion vereinfacht wird, weil die *Reduction Functions* einfacher Identitäten zwischen den Begriffen aus T_2^* und T_1 abbilden können. Schaffner (Schaffner 1969) wendet dieses Konzept beispielsweise auf den Versuch an, die klassische Genetik auf die molekulare zu reduzieren. Der Sinn und der Erfolg einer solchen Reduktion sind bis heute umstritten.

Auf dieses Konzept Schaffners scheint auch Glimcher mit seiner Vorstellung von einer Reduktion der Psychologie und Ökonomie auf die Neurowissenschaften zu bauen. An den oben dargelegten Ausführungen lässt sich erkennen, wie weit Glimcher seine Leser bereits in das Feld der Reduktion hineingeführt hat, ohne die Probleme der Reduktionisten, die auch seine Ideen beeinflussen, dargestellt zu haben. Zwar hat er Fodors bekannten Aufsatz von 1974 erwähnt, indem er kurz dessen Argument von der Psychologie als emergente, und daher nicht auf die Neurowissenschaften reduzierbare, Wissenschaft vorstellt (ein Begriff übrigens, mit dem Glimcher selbst ein paar Seiten zuvor gegen die vollständige Nagel-Reduktion argumentiert hatte). Doch die vielen weiteren Einwände, die bereits allein in demselben Aufsatz zur Sprache kamen, ignoriert Glimcher, obwohl sie auch auf sein Konzept zutreffen. Natürlich ist er nicht verpflichtet, die Wirkung seiner sorgfältig ausgearbeiteten Theorie durch die Auseinandersetzung mit allerlei Kritikpunkten zu schmälern. Doch sind die Kritikpunkte nicht neu, es ist relativ vorhersehbar, dass Gegner der Neuroökonomie ihm genau diese vorhalten werden. Dadurch, dass er selbst keine Stellung dazu nimmt, werden seine Ausführungen zu einem guten Ansatzpunkt für Kritiker. Schade, dass Gul und Pesendorfer diesen Ansatzpunkt nicht genutzt haben, denn wie gesagt, Glimchers hier zitiertes Buch erschien zwar drei Jahre nach Guls und Pesendorfers Aufsatz, die Ideen sind unter Neuroökonomien aber sicher nicht ganz plötzlich entstanden. Im Folgenden werden wir weitere von Fodors Argumenten gegen wissenschaftliche Reduktionen betrachten, die er vor allem in seinem bekanntgewordenen und von Glimcher zitierten Aufsatz von 1974 veröffentlicht hatte. Wir werden sehen, dass nicht nur die Emergenz ein Problem für Reduktionsvorhaben ist. Fodor kommt auch auf die Form und Wirksamkeit von Brückengesetzen zu sprechen und fordert eine strenge Autonomie der Psychologie von den Neurowissenschaften. Hier wird auch schon das Problem mit Inkommensurabilität

angesprochen, die von Kuhn noch stärker als grundsätzliches Hemmnis gegenüber Reduktionsansätzen jeglicher Art angesehen wird. Diese Argumentation werden wir im Anschluss an Fodors Ausführungen betrachten, wodurch wir zwar kein vollständiges, aber ein für unsere Zwecke ausreichendes Bild von den Problemen von Reduktionen generell erhalten.

Auch wenn Gul und Pesendorfer nicht weiter darauf eingegangen sind, ist die Reduktionsthematik noch immer ein Problem für Neuroökonomien, selbst wenn sie nicht so explizite Reduktionsgedanken hegen wie Glimcher. Denn es ist durchaus berechtigt, mit Gul und Pesendorfer zu fragen, wie eine Wissenschaft wie die Ökonomie, die kein neurowissenschaftliches Vokabular verwendet und für ihr traditionelles Modell auch keine *natural Kinds* aus den Kognitionswissenschaften benötigt, durch ebensolche Wissenschaften bereichert werden sollte. Wie soll die Ökonomie neurowissenschaftliche Erkenntnisse implementieren, wenn doch bereits das Vokabular nicht in ihre Umgebung passt? Als sprächen sie verschiedene Sprachen, ein Problem, auf das wir weiter unten beim Thema Inkommensurabilität noch zu sprechen kommen, können Ökonomen in der Vorstellung von Reduktionsgegnern mit neurowissenschaftlichen Theorien, salopp gesagt, schlichtweg nichts anfangen, sie passen nicht in ökonomische Theorien, wie ein falsches Puzzleteil nicht zu den anderen passt. Eine perfekte Reduktion würde diese Hindernisse theoretisch beheben, das Problem ist jedoch, dass es keine perfekte Reduktion gibt, Reduktionen funktionieren praktisch nicht. Fodor ist nur einer der bekanntesten Autoren, die diese Meinung verteidigen.

Eines der Probleme mit wissenschaftlichen Reduktionen, die Fodor (Fodor 1974) beschreibt und das auch Glimcher erwähnt, ist die Emergenz. Der Begriff sollte in der Philosophie des Geistes und in der Psychologie das Problem beschreiben, dass das menschliche Gehirn offensichtlich aus rein materiellen Bestandteilen besteht und gleichzeitig immateriell erscheinende Bewusstseinsprozesse hervorbringt. Fodor verwendet den Emergenzbegriff nun, um zu verdeutlichen, dass generell ein aus einzelnen Teilen bestehendes System Eigenschaften zeigen kann, die seine Einzelteile in separater Betrachtung nicht aufweisen. Darum macht es nach Fodor auch wenig Sinn, die Einzelteile des menschlichen Gehirns, also einzelne Neuronen und ihre Verbände zu untersuchen, um psychologische Phänomene zu erklären. Denn die Betrachtung der Einzelteile kann seiner Meinung nach nicht ihr Verhalten im kompletten System erklären. William Bechtel (Bechtel 1988, 95) erläutert das Problem anschaulich:

„It will not always be true that the properties we discover by studying the lower level entities in isolation will be the ones that are critical to explaining their performance in a complex system. For example, in studying a transistor taken out of a radio we

may not attend to those of its properties that enable it to serve its function in the radio. Similarly, although studying the properties of amino acids in isolation may reveal their primary bonding properties, it may not reveal to us those binding properties that give rise to secondary and tertiary structure when the amino acids are incorporated into protein molecules.“

Um eine ähnliche Situation könnte es sich bei der Forschung an neuronalen Strukturen handeln. Für Bechtel steht zu befürchten, dass sich Wissenschaftler nur dann derjenigen Eigenschaften neuronaler Systeme gewahr werden, die ihnen erlauben, kognitive Tätigkeiten auszuführen, wenn sie die Systeme untersuchen, *während* sie kognitive Tätigkeiten ausführen, und nicht etwa, wenn sie sie isoliert betrachten. Auf neurowissenschaftlicher Ebene wäre es demnach gar nicht möglich, diejenigen Eigenschaften des Gehirns zu untersuchen, die die Psychologie erforscht. Bei einer Reduktion würde diese Möglichkeit, die menschliche Psyche zu erforschen, verloren gehen. Die psychologischen Methoden würden durch neuroökonomische abgelöst und würden nicht mehr praktiziert werden können, alles wäre nur noch neurowissenschaftliche Forschung. Ein Weg, dieses Problem zu umgehen, könnte sein, bei einer Reduktion die reduzierende Theorie so zu verändern, dass sie die Eigenschaften von Entitäten, die dann zutage treten, wenn sie in übergeordneten Theorien eingebettet sind, selbst beinhaltet. Das würde allerdings zu der Frage führen, wie stark die untergeordnete Theorie verändert werden darf oder sollte, um eine Reduktion möglich zu machen. Sie könnte an viele Eigenschaften übergeordneter Theorien angepasst werden. Doch dann besteht die Gefahr, dass sie so stark aufgeweicht werden, dass eine Reduktion eine triviale Angelegenheit würde. Zu sehr dürfte die reduzierende Theorie also nicht modifiziert werden, aber welches Maß an Modifikation angemessen und welches zu groß ist, darüber herrscht auch unter Reduktionisten keine Einigkeit.¹⁵

Bei einer Reduktion der Psychologie auf die Neurowissenschaften entstünde also die Gefahr, dass Eigenschaften von Systemen in der Forschung nicht erkannt würden. Das würde den Erkenntnisgewinn stark behindern. Zudem bestünde ganz generell das Problem, dass Errungenschaften der übergeordneten Theorie verloren gingen, da psychologische Theorien natürliche Arten anders klassifizieren als neurowissenschaftliche Theorien, da die Psychologie andere Arten von Relationen abbilden möchte als die Neurowissenschaften und dabei anderes Vokabular verwendet (Fodor 1974, 101–106). Bechtel illustriert das Problem mit einer Analogie (Bechtel 1988, 77): Man stelle sich unsere Begriffe für Farbe (blau, rot, grün, ...)

¹⁵ Für eine eingehendere Betrachtung dieser Diskussion siehe beispielsweise Bechtel 1988, 95–96.

und Größe (groß, klein, ...) vor. Grundsätzlich korrespondiert eine Farbklassifikation nicht mit einer Größenklassifikation, das heißt, es gibt große und kleine rote Objekte und große und kleine blaue, obwohl alle farbigen Objekte auch hinsichtlich ihrer Größe beschrieben werden können. Bechtel ist der Meinung, Fodor sehe eine ähnliche Situation zwischen der psychologischen Terminologie und dem neurowissenschaftlichen Vokabular. Trifft das zu, so sind keine Brückengesetze zwischen den Vokabularien der beiden Disziplinen möglich und somit auch keine Reduktion der Gesetze der einen Theorie auf die der anderen. Die beiden Vokabularien könnten nicht über Gesetze miteinander verbunden werden, weil sie unterschiedliche Relationen und Klassifikationen abbilden, die unter Umständen nichts miteinander zu tun haben und nicht vergleichbar sind. Es wäre, als müssten Äpfel zu Birnen modelliert werden, die Vokabularien lassen sich nach Fodor einfach nicht gegenseitig übersetzen, sie sind inkommensurabel. Daher lassen sich keine Brückengesetze bilden, wodurch an eine Reduktion der Psychologie auf die Neurowissenschaften, nicht zu denken ist. Denn ohne Brückengesetze gibt es in einer heterogenen Reduktionssituation keine Reduktion.

Diese Achillesferse des Reduktionismus nutzt Fodor (Fodor 1974) weiter aus, indem er die Möglichkeit von Brückengesetzen zwischen Psychologie und Neurowissenschaften noch aus einem weiteren Grund ausschließt, der multiplen Realisierbarkeit. Dieses Argument hat die Annahme zur Basis, dass die Eigenschaften einer übergeordneten Theorie nicht immer direkt mit Eigenschaften der untergeordneten, reduzierenden Theorie übereinstimmen:

„ $S_1x \rightarrow S_2x$ “ sei ein Gesetz einer Spezialwissenschaft X, die reduziert werden soll.

„ $P_i x \rightarrow P_j^* x$ “ sei ein Gesetz einer Wissenschaft, die Wissenschaft X reduzieren soll (mit $i = 1, 2, 3, \dots, n$; $j = 1, 2, 3, \dots, m$).

Die Eigenschaften der verschiedenen Wissenschaften seien verbunden durch Brückengesetze:

- (1) „ $S_1x \Leftrightarrow P_1x \vee P_2x \dots P_nx$ “
- (2) „ $S_2x \Leftrightarrow P_1^*x \vee P_2^*x \dots P_m^*x$ “

Es lässt sich leicht sehen, dass die Eigenschaften S_1 und S_2 der speziellen Wissenschaft X nicht nur eine sondern mehrere Korrespondenten in der reduzierenden untergeordneten Wissenschaft haben. Genau hier liegt das Problem: Die Relation von P_n zu S_1 beziehungsweise P_m^* zu S_2 ist nicht eins zu eins, sondern viele zu eins. Das heißt, die Eigenschaften S_1 und S_2 sind multipel realisierbar. Fodor verwendet zur Veranschaulichung ein Schaubild, an welches das folgende Schaubild angelehnt ist (nach Fodor 1974, 109):

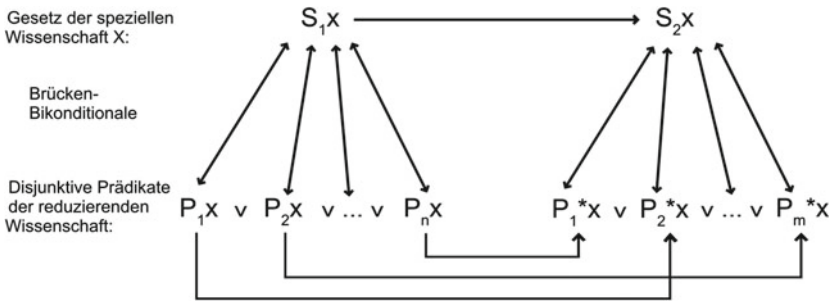


Fig. 6.1 Multiple Realisierbarkeit

Das Bild veranschaulicht, dass es in der reduzierenden Theorie mehrere Möglichkeiten gibt, die Eigenschaft S_1 der speziellen Wissenschaft abzubilden. Ebenso gibt es mehrere Möglichkeiten, die Eigenschaft S_2 in der reduzierenden Wissenschaft abzubilden. Problematisch ist das, wenn man versucht, das Gesetz der speziellen Wissenschaft „ $S_1x \rightarrow S_2x$ “ mit den Prädikaten der reduzierenden Theorie auszudrücken: Die Prädikate $P_1x \dots P_nx$ müssen über Konditionale mit den Prädikaten $P_1^*x \dots P_m^*x$ verbunden werden. Aber welche mit welchen? Ist „ $P_1x \rightarrow P_1^*x$ “ die richtige Verbindung, um das ausdrücken zu können, was „ $S_1x \rightarrow S_2x$ “ in der speziellen Wissenschaft ausgedrückt hat? Oder muss es „ $P_1x \rightarrow P_4^*x$ “ lauten? Oder doch „ $P_1x \rightarrow P_{18}^*x$ “? Nach Fodor lässt sich das nicht herausfinden. Hat etwas zu einem bestimmten Zeitpunkt die Eigenschaft P_n , dann hat es zu diesem Zeitpunkt auch die Eigenschaft S_1 . Ebenso gilt, wenn etwas zu einem bestimmten Zeitpunkt die Eigenschaft P_m^* hat, so hat es zu diesem Zeitpunkt auch die Eigenschaft S_2 . Das heißt auch, dass es n Wege gibt, in der untergeordneten Wissenschaft die Eigenschaft S_1 zu realisieren und m Wege, S_2 zu realisieren. Jede der untergeordneten Eigenschaften wird diese eine korrespondierende übergeordnete Eigenschaft realisieren. Wenn man also weiß, dass etwas die Eigenschaft S_1 besitzt, ist nicht entscheidbar, welche der n P -Eigenschaften es in der untergeordneten Wissenschaft zeigt, oder welche P^* -Eigenschaft es zeigt, wenn es Eigenschaft S_2 besitzt. Elliott Sober (Sober 1999, 545–546) liefert ein Beispiel, das in unserem Fall gut passt: Nehmen wir an, verschiedene Formen physischer Systeme können einen Geist haben. So ein Geist kann aus Neuronen gebaut sein, oder aus Silikonchips. Ein individueller Geist hat seine psychologischen Eigenschaften wegen der, und korrespondierend zu den, physiologischen Eigenschaften des Systems. Ein anderer Geist kann die gleichen psychologischen Eigenschaften haben, aber das bedeutet nicht, dass diese beiden Geister auch

physikalische Eigenschaften teilen; den gleichen psychologischen Eigenschaften könnten ganz verschiedene physische Eigenschaften zugrunde liegen. Die schiere Beobachtung der gleichen psychologischen Eigenschaften lässt keine Schlüsse auf die physische Realisierung zu. Ein etwas praktischerer Zugang zu diesem Beispiel wären beispielsweise Robo Advisors: Nehmen wir einen Bankkunden, der eine Anlage tätigen möchte. Er verwendet das Online-Banking-Portal seiner Bank und gibt in einen Fragebogen einige Angaben dazu ein, wieviel Geld er in welcher Risikoklasse für welchen Zeitraum anlegen möchte. Als Antwort von der Bank erhält er einige Vorschläge, sein Geld in einem börsengehandelten Fonds (ETF) anzulegen. Der Kunde stimmt zu und beauftragt die Bank, eine Anlage zu eröffnen, zu handeln und bei Veränderungen mit Verlegungen seiner Einlage gemäß seiner Risiko- und anderer im eingänglichen Fragebogen genannter Präferenzen zu reagieren. Die Dienstleistung, die der Kunde hier gebucht hat, kann die Bank von einem menschlichen Bankberater ausführen lassen oder von einem Robo Advisor. Für den Kunden sehen beide Möglichkeiten gleich aus, er wird nicht unterscheiden können, ob ein Mensch oder ein Computer seinen Auftrag ausführt. Wir sehen an diesem Beispiel die beiden Aspekte, die für Fodors Argument gegen den Reduktionismus wichtig sind: Erstens ist die Dienstleistung der Bank multipel realisierbar, in Fleisch und Blut, oder in Silikon. Zweitens ist von außen (für den Kunden) nicht erkennbar, welche Realisierung vorliegt, also ob seine Anlage nun von einem Roboter verwaltet wird oder von einem menschlichen Bankmitarbeiter. Nocheinmal zurück zu unserer Reduktion von „ $S_1x \rightarrow S_2x$ “: Es gibt erstens jeweils mehrere Möglichkeiten, die Eigenschaften S_1 und S_2 mithilfe von Prädikaten der reduzierenden Wissenschaft auszudrücken und es ist zweitens nicht klar, welches dieser Prädikate mit welchem verbunden werden muss, um das gleiche auszusagen wie „ $S_1x \rightarrow S_2x$ “. Fodor schließt daher als Konsequenz aus der multiplen Realisierbarkeit, dass Gesetze der übergeordneten Wissenschaft nicht durch Gesetze der untergeordneten Wissenschaft erklärt werden können. Es wird seiner Meinung nach nicht gelingen, „ $S_1x \rightarrow S_2x$ “ jemals durch die Prädikate $P_1x \dots P_nx$ und $P_1^*x \dots P_m^*x$ auszudrücken. Das kommt von der disjunktiven Form des untergeordneten Gesetzes, das entsteht, wenn die disjunktiven Prädikate der reduzierenden Wissenschaft über ein Konditional verbunden werden (vgl. Fig. 6.1):

$$(3) \text{ „}P_1x \vee P_2x \vee \dots \vee P_nx \rightarrow P_1^* \vee P_2^* \vee \dots \vee P_m^*\text{“}$$

Nach Fodor (Fodor 1974, 108–110) können Gesetze nicht disjunktiv sein. Denn (3) hat die Form eines Arguments mit den Prämissen $(P \rightarrow R)$ und $(Q \rightarrow S)$, aus denen die Konklusion $(P \vee Q) \rightarrow (R \vee S)$ folgt. Das heißt, man argumentiert

von „es ist ein Gesetz, dass P R hervorbringt“ und „es ist ein Gesetz, dass Q S hervorbringt“ zu „es ist ein Gesetz, dass (P oder Q) (R oder S) hervorbringen“. Nach Fodors Ansicht wird das aber niemand wollen, auch er nicht (Fodor 1974, 109):

„I think, for example, that it is a law that the irradiation of green plants by sunlight causes carbohydrate synthesis, and I think that it is a law that friction causes heat, but I do not think that it is a law that (either the irradiation of green plants by sunlight or friction) causes (either carbohydrate synthesis or heat).“

Daher kann (3) kein Gesetz sein. Ebenso wenig können (1) und (2) Gesetze sein. Damit sind auch Brückengesetze keine Gesetze. Das bringt uns wieder an den Punkt, dass es ohne Brückengesetze keine heterogene Reduktion gibt. Die Bedingungen für eine Reduktion können also nicht erfüllt werden.

Eine Reduktion der Psychologie auf die Neurowissenschaften ist für Fodor also nicht möglich. Sie ist von ihm auch nicht gewollt: Aus der Unmöglichkeit von Brückengesetzen und der oben besprochenen Unterschiedlichkeit der Vokabulare und Klassifikationen¹⁶ macht Fodor eine starke Autonomie der übergeordneten Wissenschaft Psychologie von den untergeordneten Neurowissenschaften aus. Aufgrund dieser Autonomie können die beiden Wissenschaften nicht voneinander lernen und müssen sich getrennt voneinander entwickeln. Damit erteilt er bereits der Idee einer Reduktion eine Absage. Für dieses Postulat strenger Autonomie erntete Fodor allerdings auch heftige Kritik, beispielsweise von Robert Richardson (Richardson 1979), der daran erinnert, dass auch Nagel kein Eins-zu-Eins-Mapping von übergeordneten Begriffen zu untergeordneten gefordert hatte, sondern dass auch das Nagel-Modell multiple Realisierungen derselben übergeordneten Eigenschaft zuließ, solange erklärt werden konnte, warum die verschiedenen untergeordneten Eigenschaften dieselbe höhergeordnete erklären (Bechtel 1988, 79). Auch Bechtel warnt, dass es gefährlich sein könnte, die Psychologie von den Erkenntnissen der Neurowissenschaften abzugrenzen und hat auch eine Vorstellung davon, wie diese hilfreich für Psychologen sein könnten (Bechtel 1988, 82):

„Neuroscience research can at least show that the operations proposed in a particular psychological theory do not correlate with processes that are performed in the brain. This provides at least *prima facie* reason for looking for an alternative psychological theory that treats a different kind of function as basic. There are even further contributions that neuroscience might make. Information about the type of processes occurring in the brain when specific cognitive activity is being performed may suggest information processing models psychologists might usefully investigate. Thus,

¹⁶ Siehe oben S. 172.

it seems plausible, if not likely, that information from neuroscience may be crucial in developing and evaluating psychological theories. To insist on the strong autonomy of psychology is to cut oneself off from such useful guidance and valuable information.“
[Hervorhebung im Original]

Wie wir uns erinnern, hatten auch Gul und Pesendorfer ganz ähnlich eine strenge Autonomie der Ökonomie von den Neurowissenschaften verlangt. Von Harrison wurden sie dafür als unnötig isolierend kritisiert. Die Frage, ob Psychologie oder Ökonomie von den Neurowissenschaften lernen können, ist eine heikle Glaubensfrage. Für Fodor oder Gul und Pesendorfer ist es so etwas strukturell unmöglich, während Leute wie Bechtel oder Harrison eher der Meinung sind, dass die Idee von der gegenseitigen Befruchtung nicht verworfen werden sollte, bevor man ihr nicht wenigstens probehalber nachgegangen ist.

Fodor jedenfalls sieht die Autonomie von speziellen Wissenschaften als unausweichlich und als guten Grund gegen eine Reduktion der Psychologie auf die Neurowissenschaften. Das würde auch die Reduktion der Ökonomie auf die Neurowissenschaften betreffen, denn zumindest in Glimchers dahingehenden Ideen spielt die Psychologie eine Rolle, wie wir gesehen haben.¹⁷ Fodors Argumente, die er sein Forscherleben lang gegen den Reduktionismus vorgebracht hat, sind berühmt geworden und haben ihre Wirkung nicht verfehlt. Aber sie sind nicht die einzigen Argumente, die gegen Theorienreduktion vorgebracht wurden. Mit einem weiteren, der Inkommensurabilität, werden wir uns noch beschäftigen, da es nicht nur in diesem Zusammenhang wichtig ist, sondern auch Guls und Pesendorfers Unverständnis für die Hinzuziehung neurowissenschaftlicher Erkenntnisse in die Ökonomie veranschaulicht.

Die Inkommensurabilität als anti-reduktionistisches Argument findet sich vor allem bei Feyerabend (Feyerabend 1962) und Kuhn (Kuhn 1962). Wir werden uns im Folgenden vor allem auf Kuhns Konzeption von Inkommensurabilität konzentrieren, da sie Aspekte umfasst, denen wir weiter oben begegnet sind, als es darum ging, dass Behavioristen und Neuroökonomien grundsätzlich zu keiner Einigung gelangen können.¹⁸

Kuhn hat eine ganz andere Vorstellung von Theoriensukzession als durch Reduktion. Er ist vermutlich der bekannteste Vertreter der Idee, dass das Verhältnis zweier Theorien nicht Reduktion ist, sondern *Replacement*. Dabei wird eine Theorie, nach Phasen der Normalwissenschaft, über wissenschaftliche Revolutionen durch eine andere abgelöst. Kuhn verwendet den Begriff Inkommensurabilität vor allem zur Charakterisierung des Verhältnisses zweier solcher Theorien, er

¹⁷ Siehe oben S. 170.

¹⁸ Siehe oben S. 164–165.

stammt aber bereits aus der griechischen Antike, wie Sharrock und Read (Sharrock et al. 2002, 141–142) erläutern und wurde dort offenbar vor allem von den Pythagoräern verwendet (Sharrock et al. 2002, 141): „Literally, incommensurability means the impossibility or unavailability of a common system of measure.“ Dinge, die inkommensurabel sind, können also nicht mit gleichem Maß gemessen werden. Die Mathematiker um Pythagoras verwendeten den Begriff für das Verhältnis der Längen der Katheten und der Hypothenuse eines rechtwinkligen Dreiecks. Da die eine eine rationale, die andere eine irrationale Zahl ist, konnten sie nicht mit gleichem Maß gemessen werden, waren also inkommensurabel. Bei Kuhn bedeutet die Inkommensurabilität insofern die Abwesenheit eines gemeinsamen Maßes, dass zwischen den beiden Theorien keine direkten Punkt-für-Punkt Vergleiche möglich sind, was aber nicht bedeutet, dass überhaupt keine Vergleiche möglich sind. Die Inkommensurabilität besteht nach Hoyningen–Huene (Hoyningen–Huene 1989) aus drei verschiedenen Aspekten, von denen vor allem der zweite und dritte für unseren Kontext interessant sind. Der erste Aspekt betrifft Definitionen und Normen der von einer Revolution betroffenen Wissenschaften selbst. Hoyningen–Huene (Hoyningen–Huene 1989, 203) erläutert:

„Mit einer wissenschaftlichen Revolution ändern sich der Bereich der notwendigerweise als auch der Bereich der legitimerweise zu bearbeitenden wissenschaftlichen Probleme. Probleme, deren Beantwortung für die ältere Tradition von zentraler Bedeutung war, können als veraltet oder unwissenschaftlich verschwinden; Fragen, die für die ältere Tradition nicht existierten oder deren Beantwortung trivial war, können große Wichtigkeit erlangen.“

Das ist ein interessanter Punkt, für uns aber nicht ganz so relevant wie die nächsten beiden Punkte, weshalb wir dieses Thema an dieser Stelle hinter uns lassen und gleich zu Aspekt 2 der Inkommensurabilität weitergehen. Dieser zweite Aspekt bezieht sich auf die in der Wissenschaft verwendeten Begriffe und deren Verwendungsweisen. Hoyningen–Huene (Hoyningen–Huene 1989, 204):

„Nach einer Revolution werden viele zur vorangehenden normalwissenschaftlichen Tradition gehörenden Verfahrensweisen und Begriffe zwar weiterhin, aber auf modifizierte Weise verwendet. Besondere Wichtigkeit kommt dabei [...] der Änderung von Begriffen bzw. der Verwendungsweise von Begriffen zu, der sogenannten Begriffsverschiebung [meaning change].“

Da neue Paradigmen aus alten entstehen, wird vieles an Vokabular und auch Laborausüstung in die neue Theorie übernommen. Die von Hoyningen–Huene

erwähnte Begriffsverschiebung hat zur Folge, dass alte Ausdrücke zwar noch verwendet werden, jedoch auf andere Weise und in einem anderen Zusammenhang. Die Sprache der Wissenschaftler hängt davon ab, unter welchem Paradigma sie arbeiten. Hoyningen-Huene identifiziert eine extensionale und eine intensionale Seite der Begriffsverschiebung. Die extensionale beinhaltet, dass Objekte aus der Extension eines Begriffs in die Extension eines anderen Begriffs übergehen. Dabei sind die Extensionen der Begriffe verschieden. Kuhn (Kuhn 1962, 128–129) gibt dazu das Beispiel eines Astronomen, der vom ptolemäischen zum kopernikanischen Weltbild wechselt und nun die Erde als Planeten bezeichnet, den Mond oder die Sonne, die früher als Planeten bezeichnet wurden, dagegen nicht mehr. Der intensionale Aspekt der Begriffsverschiebung beinhaltet, dass sich die Bedeutung der Begriffe ändert, da sich die Eigenschaften der Objekte ändern, die mit dem Begriff bezeichnet werden. Das betrifft aber nur solche Begriffe, bei denen sich Eigenschaften ändern, die in der Definition des Begriffs vorkommen. Hoyningen-Huene (Hoyningen-Huene 1989, 206) ist allerdings der Meinung, dass diese Forderung Kuhns nicht zu erfüllen ist. Seiner Meinung nach müssten auch bestimmte Änderungen von Eigenschaften von Extensionselementen eine Änderung des Begriffs nach sich ziehen (das ist für uns momentan nicht von besonderer Wichtigkeit, weshalb wir nicht weiter darauf eingehen).

Der dritte Aspekt der Inkommensurabilität ist für Kuhn auch der fundamentalste. Er besagt, dass die Vertreter konkurrierender Theorien in verschiedenen Welten arbeiten. Zwar arbeiten die Wissenschaftler während einer Revolution mit den gleichen Instrumenten wie in der Zeit der normalen Wissenschaft, sehen jedoch bei Ihren Experimenten neue und andere Dinge. Eines von Kuhns Beispielen (Kuhn 1962, 118–119) sind ein Aristoteliker und Galilei, die beide einen schweren Körper an einer Kette hin- und herschwingen sehen. Der Aristoteliker sieht darin einen Körper, der gehemmt fällt. Da er von der Kette gehalten wird, kann er erst nach einiger Zeit und mühsamen Bewegungen am niedrigsten Punkt zur Ruhe kommen. Galilei dagegen sieht in dem schwingenden Körper ein Pendel, dem es beinahe gelingt, die immer gleiche Bewegung unendlich auszuführen. Die beiden Vertreter unterschiedlicher Paradigmen schauen aus dem gleichen Blickwinkel auf den gleichen Vorgang und sehen doch etwas anderes. Kuhn (Kuhn 1962, 111) spricht sogar recht drastisch über diese Veränderung des Wahrgenommenen:

„It is rather as if the professional community had been suddenly transported to another planet where familiar objects are seen in a different light and are joined by unfamiliar ones as well. Of course, nothing quite that sort does occur: there is no geographical transplantation; outside the laboratory everyday affairs usually continue as before. Nevertheless, paradigm changes do cause scientists to see the world of their research-engagement differently.“

Zusammengefasst ändern sich bei einem Paradigmenwechsel die Bedeutungen von Begriffen und Beobachtungen. Mehr noch: Mit einer Theorie kommt auch ein ganzes Bündel an Wissen und Begriffen, im Grunde eine ganze Weltsicht. Bei einem Theorienwechsel verändern sich auch sie, bis zu dem Maß, dass die Forschenden das Gefühl bekommen, sich in einer ganz anderen Welt wiederzufinden. Sie verwenden zwar dieselbe Laborausrüstung wie zuvor, doch während und nach einem Paradigmenwechsel sehen sie damit neue und andere Dinge. Das bedeutet, und das ist hier ganz wichtig, eine Abhängigkeit der Beobachtungen, die Wissenschaftler bei ihrer Arbeit machen und die in die Gesetze einfließen, die sie formulieren, von dem Paradigma, unter dem sie forschen. Beobachtungen und Begriffe sind nicht neutral bei Kuhn, sie sind theoriengeladen. Diese Theoriegeladenheit macht die Inkommensurabilität zu einem interessanten Argument gegen Reduktionismus. Denn wie können Theorien, die offenbar so verschieden sind, reduziert werden? Wie können Identitäten zwischen Begriffen gefunden werden, wenn sich ihre Bedeutung verändert? Dieses Problem ist besonders schwerwiegend, da Wissenschaftler die Änderung in den Bedeutungen von Begriffen laut Kuhn nicht einmal unbedingt bemerken müssen. Sie können annehmen, dass die Sprache, die ein Kollege aus einer anderen Theorie spricht, dieselbe ist, die sie auch sprechen. Es ist gut möglich, dass sie Begriffe mit unterschiedlichen Bedeutungen verwenden und deshalb das, was ihr Gegenüber sagt, eigentlich falsch verstehen. Folglich könnten die Forscher, die Brückengesetze formulieren, ebenso Begriffe und Bedeutungen falsch verstehen, auch ohne es zu bemerken, und dadurch falsche Identitäten zwischen den Theorien aufstellen. Das führt zu einer fehlerhaft oder überhaupt nicht funktionierenden Reduktion. Solche Missverständnisse können, wenn sie unbemerkt bleiben, heillose Konsequenzen haben, wenn Forscher mit den so entstandenen neuen Theorien arbeiten, weil sie auf falschen Annahmen beruhen. Unter solchen Umständen wären Theorienreduktionen fahrlässige Unternehmen. Daher ist die Inkommensurabilität verschiedener Wissenschaften ein starkes Argument gegen Reduktionismus, auch Fodor bezieht sich auf eine solche Idee bei seinem oben¹⁹ besprochenen Argument von verschiedenen Klassifikationssystemen und Vokabularen zwischen den Wissenschaften.

An dieser Stelle wird auch klar, warum Gul und Pesendorfer mit vollkommenem Unverständnis darauf reagieren, dass manche ihrer Kollegen neurowissenschaftliche Erkenntnisse in ökonomische Modelle aufnehmen möchten. Sie arbeiten unter einem anderen Paradigma, in dem Erkenntnisquellen, wie sie in den kognitiven Neurowissenschaften üblich sind, nicht nur keine Rolle spielen,

¹⁹ Siehe oben S. 174–175.

sondern auch keine Rolle spielen dürfen. Ihre Vorstellung von der Ökonomie als Wissenschaft ist inkommensurabel mit der Vorstellung der Verhaltensökonomie. Solange keine der beiden Seiten ihr Paradigma verlässt und zum jeweils anderen wechselt, werden sich die Probleme zwischen den verschiedenen Ökonomen nicht lösen lassen. Laut Kuhn können sie ja nicht einmal verständlich und verlässlich darüber diskutieren. Die Ökonomen werden also weiterhin uneins sein. Laut Kuhn ist das aber nicht schlimm, für ihn ist das eher eine Art natürlicher Lauf der wissenschaftlichen Forschung, dem schon viele wissenschaftliche Gemeinschaften ausgesetzt waren und noch viele ausgesetzt sein werden.

Wie man sieht, stammen die besprochenen Argumente bereits aus den 50er bis 70er Jahren und sind durchaus gründlich diskutiert worden. Bisher hat sich noch keine Lösung darüber aufgetan, ob die Psychologie, deren damalige Lage man in diesem Fall durchaus mit der heutigen Lage der Verhaltensökonomie vergleichen kann, auf die Neurowissenschaften reduziert werden kann oder sollte. Inzwischen wird wissenschaftlicher Reduktionismus von den meisten Wissenschaftstheoretikern als eher metaphysische Position betrachtet. Neuroökonomien sollten sich aber durchaus mit verschiedenen Konzeptionen des Reduktionismus und ihren Kritiken auseinandersetzen, denn sie müssen sich zurecht die Frage gefallen lassen, wie Erkenntnisse aus neuronalen Datenquellen Eingang in nicht neuronal basierte ökonomische Modelle finden sollen. Wie wir gesehen haben, scheint Glimcher eine durchaus differenzierte Vorstellung von partieller Theorienreduktion zwischen den beiden Wissenschaften zu haben. Wie wir aber auch gesehen haben, treffen auch auf seine Vorstellung diejenigen Kritiken zu, die bereits vor Jahrzehnten Reduktionskonzeptionen entgegengebracht wurden und die dazu geführt haben, dass Reduktionismus heute als eine metaphysische Position angesehen wird. Wie Gul und Pesendorfer sagten²⁰, das *Vehicle* zwischen den beiden Wissenschaften fehlt, und meine Prognose ist, dass auch Glimcher dafür keine Lösung finden wird. Das wird Neuroökonomien nicht von ihrer Arbeit abhalten, jedoch sollten sie sich der Hintergründe und Problematiken bewusst sein, wenn sie versuchen, Erkenntnisse aus der einen Wissenschaft für Erkenntnisse in der anderen zu nutzen. Naivität im Umgang mit den Problemen, die in diesem Unterkapitel besprochen wurden, könnte zu unsauberer Modellen und damit zum Scheitern der Neuroökonomie führen.

²⁰ Siehe oben S. 161.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.



Nachdem wir uns einige Kritikpunkte und Probleme der Neuroökonomie angeschaut haben, bleibt die Frage, wie sich das Verhältnis von Ökonomen und Neurowissenschaftlern entwickeln wird. Denn davon hängt die künftige Entwicklung der Neuroökonomie als Wissenschaft ab. Sind die Differenzen zu groß, die gegenseitig vorgebrachten Kritiken unüberwindbar, wird eine langfristig stabile und fruchtbare Kooperation kaum möglich sein.

Werfen wir also einen Blick auf zukünftige Entwicklungsmöglichkeiten für die neuroökonomische Zusammenarbeit. Welche Probleme werden noch in Zukunft eine Rolle spielen und welche Entwicklungschancen hat die Neuroökonomie? Welche Probleme sind lösbar und welche unlösbar? In Kapitel 6 zu den Kritiken an der Neuroökonomie haben wir die vorgebrachten Kritiken aufgeteilt in grundsätzliche Kritiken, die eine Zusammenarbeit immer beeinträchtigen könnten und solche Kritiken, die grundsätzlich lösbar sind. Ähnlich verhält es sich mit den Problemen, die wir nun als Abschluss betrachten, da sie aus den oben angesprochenen Kritiken erwachsen. Manche sind generell lösbar und sollten gelöst werden. Andere dagegen lassen sich grundsätzlich nicht lösen, sodass bestenfalls ein Weg gefunden werden muss, mit ihnen umzugehen. Beginnen wir mit den lösbaren Problemen:

Die lösbaren Probleme sind zwar grundsätzlich lösbar, das ist ein großer Vorteil gegenüber unlösbaren Problemen, aber wenn sie nicht gelöst werden, können sie erheblichen Schaden für die Wissenschaft Neuroökonomie anrichten. Sie können Wissenschaftler abschrecken und zur Ablehnung der Neuroökonomie führen. Damit sind solche Wissenschaftler gemeint, die als Forschende für die Neuroökonomie wertvoll sind oder sein könnten, vor allem diejenigen, die aus den Mutterwissenschaften stammen. Besonders als noch relativ junge Wissenschaft braucht die Neuroökonomie kreative und offene Nachwuchskräfte, die die Forschung betreiben und in die Zukunft tragen. Gemeint sind

aber auch solche Wissenschaftler, die nicht direkt an der neuroökonomischen Forschung beteiligt sind, aber ihre Erkenntnisse für ihre eigene Forschung nutzen könnten, wie beispielsweise aus nicht-kognitiven Bereichen der Psychologie. Diese Verbreitung der Erkenntnisse trüge zur Festigung der Relevanz der neuroökonomischen Forschung bei. Nicht zuletzt um beide Arten von Akzeptanz unter Wissenschaftlern nicht zu gefährden, sollten die lösbaren Probleme auch gelöst werden. Die Probleme stammen aus den prozeduralen Kritiken, die wir oben in Abschnitt 6.1 diskutiert haben. Das waren erstens die überzogenen Versprechungen von Neuroökonomern bezüglich der Leistungsfähigkeit der Neuroökonomie, zweitens die Kritiken an der unzureichenden Technik, die für kognitiv-neurowissenschaftliche Forschung verwendet wird, drittens die Kritik an schlechten neurowissenschaftlichen Forschungsstandards, wie zu geringe Probandenzahlen, viertens ein mangelhaftes Forschungsdatenmanagement, das die Nachprüfbarkeit von Ergebnissen verringert und fünftens die augenscheinlich arglose Verwendung von *reverse Inference* als gängige Schluss-Praxis.¹

Aus den überzogenen Versprechungen wächst leicht das Problem geringer Glaubwürdigkeit, was sehr problematisch sein kann, aber auch relativ leicht lösbar. Wie wir in Abschnitt 6.1.1 gesehen haben, scheinen die Neuroökonomern das auch bemerkt zu haben, da jüngere Aussagen über das Potenzial der Neuroökonomie weniger überzogen ausfallen als die kritisierten älteren. Wenn die Neuroökonomern das weiterhin so halten, können sie hoffentlich ihre „marketing bloopers“ aus Anfangszeiten, wie Harrison es nannte (Harrison 2008a, 304), überwinden.

Das zweite Problem mit der unzureichenden Technik wird sich, wie in Abschnitt 6.1.2 erörtert, durch den zu erwartenden Fortschritt der Instrumententechnik in naher Zukunft zum großen Teil erledigen. Viel drängender ist das dritte Problem, das die angemessene Zahl der Probanden pro Studie betrifft. Dieses Problem muss gelöst werden, nicht nur um die Neuroökonomie als Wissenschaft etablieren zu können, sondern auch, weil dies eine Frage ist, die ebenso die Mutterwissenschaften selbst betrifft. Die Richtlinien für die Erstellung und Durchführung von fMRT-Studien (um diese geht es in dieser Kritik ja vor allem) kommen aus den kognitiven Neurowissenschaften, wo diese Art der (neuroökonomischen) Forschung durchgeführt wird. Es ist nicht nur im Interesse der Neuroökonomern, sondern der kognitiven Neurowissenschaftler genauso

¹ Da wir diese Kritiken in Abschnitt 6.1 bereits ausführlich betrachtet haben, möchte ich im Folgenden nicht noch einmal länger darauf eingehen, eine kurze Beschreibung soll hier genügen.

wie Psychologen die Frage angemessener Probandenzahlen zu klären. Die Replikationskrise hat gezeigt, dass solche Fragen eine erhebliche und vielleicht bisher unterschätzte Rolle spielen. Reliabilität ist für Forschungsergebnisse essenziell, da sind sich die Mitglieder einer jeden wissenschaftlichen Gemeinschaft sicherlich einig. Fehlt sie oder steht sie auch nur gering in Zweifel, kann das mittel- und langfristig erhebliche Probleme nach sich ziehen, sowohl innerhalb einer betroffenen Wissenschaft, wenn Erkenntnisse, auf denen neue Erkenntnisse aufgebaut werden nicht verlässlich sind, denn dann stehen auch darauf aufbauende Erkenntnisse auf unsicherem Grund; aber auch außerhalb einer Wissenschaft, wenn der Ruf nicht nur einzelner Forschender, sondern einer ganzen Gemeinschaft leidet. Es ist also keine allein neuroökonomische Aufgabe, das Problem angemessener wissenschaftlicher Standards anzugehen, sondern eine Aufgabe auch und vor allem für die kognitiven Neurowissenschaften, wollen sie in Zukunft glaubwürdig sein, sei es als Person oder weiterverwertbare Forschungsergebnisse.

Helfen könnte dabei eventuell die Lösung zu Problem vier, dem schlechten Forschungsdatenmanagement. Die Nachvollziehbarkeit von Forschungsergebnissen könnte höher sein, wenn mit den Textpublikationen zusammen mehr Forschungsdaten veröffentlicht würden, auf denen die Schlussfolgerungen und Erkenntnisse beruhen. Diese Form von Open Data ist in manchen Wissenschaften beziehungsweise Zeitschriften bereits verbreiteter als in anderen, gewinnt aber gerade an Momentum, da Forschungsförderer wie die Europäische Kommission die gleichzeitige Open-Access-Veröffentlichung von Text und den zugrundeliegenden Daten in ihrer Förderlinie Horizon 2020 zum Grundsatz gemacht haben.² Das heißt, dass die Daten, die den Textpublikationen zugrunde liegen, in einem Repository open access zugänglich gemacht werden müssen. Eine solche Praxis hilft dabei, die Ergebnisse und Schlussfolgerungen überprüfbar zu machen. Natürlich können nicht alle Arten von Daten öffentlich zugänglich gemacht werden und auch bei der EU-Kommission gibt es in begründeten Fällen die Möglichkeit zu einem Opt-Out, beispielsweise wenn Patientendaten nicht stark genug anonymisierbar sind. Aber die grundsätzliche Einstellung, offen mit seinen Daten umzugehen, kann die Überprüfbarkeit und Glaubwürdigkeit von Forschungsergebnissen erhöhen. Dadurch ließe sich nicht nur dem Vorwurf wenig nachvollziehbarer Studienergebnisse begegnen, es könnte, wie gesagt, auch helfen, durch erhöhte Transparenz verlässliche wissenschaftliche Standards durchzusetzen. Durch die Mandate von Forschungsförderern wird der offene und

² Siehe dazu H2020 Annotated Model Grant Agreement, Artikel 29.3, abrufbar unter: https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf [abgerufen am 20.07.2020]

standardisierte Umgang mit Daten immer mehr in der Praxis der täglichen wissenschaftlichen Arbeit ankommen, wodurch sich für diese Probleme eine potentiell positive Prognose geben lässt.

Es bleibt noch das fünfte Problem mit der Verwendung von *reverse Inference* als gängige Schlusspraxis. Auch dieses ist kein spezifisch neuroökonomisches Problem, sondern betrifft die Mutterwissenschaften allgemein. Auch in diesem Fall liegt es im eigenen Interesse der Mutterwissenschaften, eine Art gute wissenschaftliche Praxis für den akzeptablen Umgang mit dieser eventuell kurzfristig hilfreichen aber tatsächlich logisch falschen Schlussform zu finden. Poldrack hat sicher Recht mit seiner Meinung, dass *reverse Inference* dazu verwendet werden kann, neue Hypothesen zu formulieren, die als Basis für weitere Forschung dienen können (s. o. Abschn. 6.1.5). Er hat aber auch Recht damit, dass die Verwendung von *reverse Inference* als Schlussystem zum Erkenntnisgewinn gefährlich ist, denn die Gefahr, statt einer Abduktion einen Fehlschluss zu ziehen, ist durchaus gegeben. Sicherlich sind neue Erkenntnisse und Schlüsse in Publikationen spannender zu lesen als Vorschläge für weitere Forschungsfragen, und vor allem kognitiv neurowissenschaftliche Forschung scheint in den vergangenen Jahren ein wachsendes Forschungsgebiet zu sein, aber weder die Neuroökonomie noch die kognitiven Neurowissenschaften selbst sollten sich mit spektakulären aber schlecht replizierbaren Erkenntnissen einen Namen machen. Denn wie das Beispiel der überzogenen Versprechungen aus der Anfangszeit der Neuroökonomie zeigt, kann ein solches Vorgehen zwar viel Aufmerksamkeit erzeugen, aber auch aus den falschen Gründen. Gewiss möchte jeder Forschende eine aufsehenerregende Publikation in einer renommierten Zeitschrift, doch methodisch schlechte Arbeit bleibt methodisch schlechte Arbeit, auch wenn die Zitationszahlen gut aussehen. Nur diesen hinterherzujagen, auf Kosten guter wissenschaftlicher Standards, das hat die Replikationskrise in der Psychologie deutlich gemacht, wird früher oder später erheblichen Schaden an einer Wissenschaft verursachen und das kann sich gerade eine junge Wissenschaft wie die Neuroökonomie nicht leisten. Natürlich ist das auch ein strukturelles Problem nicht nur der neuroökonomischen Mutterwissenschaften und natürlich wäre es naiv anzunehmen, es ließe sich über Nacht lösen. Mit den gleichen Schwächen in die nächste Replikationskrise zu steuern, kann allerdings auch nicht die Basis für wissenschaftliche Arbeit sein. Dazu braucht es aber ein generelles Umdenken unter den Mitgliedern wissenschaftlicher Gemeinschaften, von den Forschern im Labor, über Berufungskommissionen, über Herausgebergremien bis hin zu Forschungsförderern, und das kann Generationen dauern.

Die besprochenen Probleme können das Verhältnis zwischen Neuroökonomien verschiedener Mutterwissenschaften untereinander sowie das Verhältnis zu anderen Wissenschaften außerhalb der Neuroökonomie negativ beeinträchtigen. Aber sie *können* grundsätzlich gelöst werden, und die wissenschaftlichen Gemeinschaften haben viel zur Lösung dieser Probleme selbst in der Hand. Das Verhältnis wird jedoch auch von systematischen Problemen beeinflusst, die die Wissenschaftler nicht selbst in der Hand haben und die sich auch nicht lösen lassen. Mit diesen muss die Neuroökonomie umzugehen lernen, wenn sie langfristig eine Rolle spielen möchte. Damit sind erstens diejenigen systematischen Probleme gemeint, die vor allem in den Kritiken von Gul und Pesendorfer zur Sprache kommen, die wir in Abschnitt 6.2 analysiert haben. Zweitens sind damit diejenigen Probleme gemeint, die sich aus wissenschaftstheoretischer Sicht ergeben, die in Abschnitt 6.3 besprochen wurden.³ Gerade weil sich diese Probleme nicht lösen lassen, werden sie das Verhältnis zwischen Ökonomen und Neurowissenschaftlern weiterhin bestimmen. Es lohnt sich also ein kurzer, zusammenfassender Blick zurück: Gul und Pesendorfers Kritiken basierten auf strukturellen, grundlegenden Problemen, die sie bei einer Zusammenarbeit zwischen Ökonomen und Neurowissenschaftlern sehen. Sie entstehen zum Beispiel dort, wo ihrer Meinung nach Neuroökonomien den Ansatz ökonomischer Arbeit nicht verstanden haben, weshalb sie ökonomische Erkenntnisse als normative Handlungsanweisungen verwenden wollen. Zudem sehen Gul und Pesendorfer die Neuroökonomie als irrelevant für die Ökonomie an, da sie keine ökonomischen Probleme anspricht. Nicht nur liefern die Erkenntnisse aus der Neuroökonomie keinen Mehrwert für die Ökonomie, die unterschiedlichen Modelle der beiden Wissenschaften werden von den Autoren auch nicht als konkurrierend angesehen, denn die ökonomische Definition von Rationalität macht keine Aussagen über physiologische Hirnprozesse, weshalb das ökonomische Standardmodell laut Gul und Pesendorfer auch nicht mit physiologischen Gehirndaten widerlegt werden kann. Ihrer Ansicht nach vermischen Neuroökonomien Mittel und Ziele der Ökonomen mit denen der Neurowissenschaftler, haben jedoch keine Möglichkeit, die in der Ökonomie beschriebenen Prozesse im Gehirn wiederzufinden, oder ökonomische Probleme mit neurowissenschaftlichen Mitteln anzusprechen. Insgesamt sind es sehr schwerwiegende, tiefgehende Kritiken, die Gul und Pesendorfer vorgebracht haben. Sie sind anders als die Kritiken an niedrigen Probandenzahlen oder geringen Publikationsraten von Rohdaten. Es sind systematische Kritiken, die sich

³ Auch an dieser Stelle werde ich die Argumente nicht noch einmal ausführlich betrachten, da sie in den entsprechenden Kapiteln bereits behandelt worden sind, eine kurze Beschreibung soll auch hier genügen.

nicht durch (ohnehin angeratene) Änderungen in der wissenschaftlichen Praxis verringern oder aus der Welt schaffen lassen. Bei diesen Kritiken geht es um ganz verschiedene Auffassungen davon, was die Ökonomie als Wissenschaft tut und tun sollte, und um die grundlegende Frage, ob die Wissenschaften Ökonomie, Neurowissenschaften und Psychologie überhaupt zusammenarbeiten können.

Im Einzelnen haben wir gesehen, dass Gul und Pesendorfers Kritiken wissenschaftstheoretische Probleme zugrunde liegen, die teilweise seit Jahrzehnten diskutiert werden. Die beiden Ökonomen kratzen mit ihren Ausführungen lediglich an deren Oberfläche, wie in Abschnitt 6.3 diskutiert wurde. Da ist einerseits das Paket an Problemen, das mit Reduktionismus einhergeht, vor allem wie ihn Glimcher beschreibt (s. o. Abschn. 6.3), was auf eine partielle heterogene Theorienreduktion hinausläuft, und andererseits das Paket an Problemen, das mit Inkommensurabilität verbunden ist, wobei Inkommensurabilität selbst auch ein Teil des ersten Pakets ist. Zu den Problemen des Reduktionismus, die wir in Abschnitt 6.3 betrachtet hatten, gehören die Kritik Feyerabends, dass nicht Theoriengesetze reduziert werden, sondern lediglich etwas, das so ähnlich aussieht. Schaffners Antwort, Theorien zu modifizieren, bevor sie reduziert werden, entkräftet diese Kritik nicht vollständig (für Details siehe Abschn. 6.3). Auch ein weiteres Problem, die Emergenz, bleibt bislang ungelöst. Diese Idee besagt, dass ein aus Einzelteilen bestehendes System Eigenschaften aufweisen kann, die seine einzelnen Teile in separater Betrachtung nicht aufweisen. Von Fodor auf die kognitiven Neurowissenschaften und Psychologie angewandt, bedeutet das, dass ein menschliches Gehirn aus rein materiellen Teilen besteht, und gleichzeitig immateriell erscheinende Bewusstseinsprozesse aufweist, die sich aus der Betrachtung der Neuronen allein nicht ergeben (s. Abschn. 6.3). Nach einer Reduktion der Psychologie auf die Neurowissenschaften wären daher diese Bewusstseinsprozesse im schlimmsten Falle nicht mehr untersuchbar, wodurch eine Menge psychologischen Erkenntnisgewinns verloren ginge. Ein weiteres von Fodor vorgebrachtes Argument gegen Reduktionismus ist die multiple Realisierbarkeit (s. Abschn. 6.3), die zur Folge hat, dass es keine Brückengesetze geben kann, da solche Gebilde keine Gesetze sind oder sein können. Die Annahme ist hier, dass die Eigenschaften einer übergeordneten Theorie nicht immer direkt mit Eigenschaften der untergeordneten Theorie übereinstimmen, sondern von mehreren Eigenschaften der untergeordneten Theorie realisiert werden können. Dabei ist das Problem, dass laut Fodor nicht klar sein kann, welche der untergeordneten Eigenschaften die übergeordnete Eigenschaft in einem bestimmten Gesetz der übergeordneten Theorie realisiert. Daher müsste ein untergeordnetes Gesetz, das denselben Zusammenhang darstellen soll wie das zu reduzierende übergeordnete

Gesetz, eine disjunktive Form besitzen, in der alle Möglichkeiten aufgezählt werden. Das Gleiche gilt für Brückengesetze. Nach Fodor können Gesetze jedoch nicht disjunktiv sein, weshalb auch Brückengesetze keine Gesetze sein können. Ohne Brückengesetze gibt es aber keine heterogenen Reduktionen, wie sie sich Glimcher vorstellt.

Die zweite Art wissenschaftstheoretischer Probleme, an deren Oberfläche Gul und Pesendorfer kratzen, ist die Inkommensurabilität. Sie ist einerseits Teil der antireduktionistischen Argumentation, andererseits auch eine eigene Quelle für Kritiken an der neuroökonomischen Zusammenarbeit, wie wir oben gesehen haben (s. Abschn. 6.3). Ein in diesem Kontext wichtiger Aspekt der Inkommensurabilität ist die Begriffsverschiebung, nach der bei einem Paradigmenwechsel Objekte aus der Extension eines Begriffs in die Extension eines anderen übergehen, wobei die Extensionen der Begriffe verschieden sind. Der zweite wichtige Aspekt der Inkommensurabilität besagt, dass die Wissenschaftler unterschiedlicher Paradigmen auch in unterschiedlichen Welten arbeiten. Sie betrachten noch immer dieselben Objekte mit denselben Instrumenten, sehen jedoch Unterschiedliches. Ein Beispiel waren der Aristoteliker und Galilei, die beide einen schweren Körper an einer Kette hin- und herschwingen sehen. Doch während Aristoteles einen gehemmt fallenden Körper sieht, sieht Galilei ein Pendel, das die immer gleiche Bewegung beinahe unendlich ausführt. Diese Veränderung des Wahrgenommenen beschreibt Kuhn sogar recht drastisch, die Forschenden können das Gefühl bekommen, sie arbeiteten nach einer Revolution in einer völlig anderen Welt. Das bedeutet, dass das, was die Wissenschaftler sehen, vom Paradigma abhängt, unter dem sie forschen und damit auch die Modelle und Gesetze, die sie formulieren. Diese Theoriegeladenheit macht die Inkommensurabilität zu einem wichtigen Argument gegen eine Reduktion, da sie es unmöglich macht, verlässliche Interpretationen eines anderen Paradigmas zu liefern oder Brückengesetze zu formulieren.

Die Inkommensurabilität eröffnet auch einen anderen Blickwinkel auf Kritiken, wie sie Standardökonominnen wie Gul und Pesendorfer äußern: In ihren Augen ist die Ökonomie als Wissenschaft frei von neurowissenschaftlichen Erkenntnissen und Einflüssen, da diese im Standardmodell keine Rolle spielen. Das ist eine vollkommen andere Vorstellung davon, wie Ökonomie funktioniert und funktionieren soll, als die der Verhaltensökonominnen. Die Ansichten von Gul und Pesendorfer sind inkommensurabel mit den Ansichten von Verhaltensökonominnen. Das macht die Kritiken von Standardökonominnen an der Neuroökonomie verständlicher, macht aber auch deutlich, dass sie grundsätzlich nicht aus dem Weg geräumt werden können, es sei denn, alle Ökonomen wechseln unter

dasselbe Paradigma. Solange das nicht geschieht, können die gerade besprochenen systematischen Kritiken nicht gelöst werden. Egal, wie viele Argumente oder scheinbar überzeugende Studien jede Seite vorbringt (solange das nicht gerade Forschende überzeugt, zum Paradigma der Verhaltensökonomien zu wechseln), wird das keinen Einfluss auf die Meinung der Standardökonomien zur Neuroökonomie haben. Das Paradigma der Verhaltensökonomien bleibt immer inkommensurabel mit dem der Standardökonomien.

Für die Frage, wie sich das Verhältnis zwischen Ökonomen und Neurowissenschaftlern entwickeln wird, heißt das also, dass es in absehbarer Zukunft immer Kritiker geben wird, die eine solche Form der Kooperation ablehnen und dass diesen Kritiken mit den besten Argumenten nicht begegnet werden kann. Das muss für die Neuroökonomien nicht so frustrierend sein, wie es vielleicht klingen mag, denn diese Einsicht bedeutet nicht nur, dass man sich mit solcher Art Kritiken für immer auseinandersetzen muss, sie bedeutet auch, dass man sie als das ansehen kann, was sie sind: grundsätzlich nicht lösbar, weil inkommensurabel mit unserer Vorstellung von der Ökonomie. Diese Kritiken sind nicht das Ergebnis schlechter wissenschaftlicher Arbeit der Neuroökonomien oder schlechter Konzeption der Neuroökonomie an sich. Dadurch entsteht auch kein Druck, die Neuroökonomie oder die neuroökonomische Arbeit zu verändern, um Kritiken zu begegnen, das wäre alles nutzlos. Das heißt nicht, dass Neuroökonomien Kritiken von Standardökonomien ignorieren sollten, aber sie können sie als inkommensurabel erkennen.

Das gilt allerdings nur für die systematischen Kritiken, die daraus entstehen, dass Kritiker und Neuroökonomien unter verschiedenen Paradigmen arbeiten. Es gilt nicht für die grundsätzlich lösbaren Kritiken, die wir oben gerade noch einmal betrachtet haben. Sie bauen in der Tat auf Zuständen in der Neuroökonomie, die sich nicht nur ändern lassen, sondern, wie wir gesehen haben, auch geändert werden sollten, auch im Sinne der Neurowissenschaften. Damit meine ich vor allem die Reproduzierbarkeit von Studienergebnissen und die Probandenzahlen. Für sie gibt es Hoffnung, weil sie nicht auf der Inkommensurabilität zweier Paradigmen basieren. Sie zu lösen ist notwendig, um nicht diejenigen Wissenschaftler abzuschrecken, die grundsätzlich Befürworter der Neuroökonomie sein könnten. Die Standardökonomien, die nicht von ihrem Paradigma abrücken, werden aus den dargelegten Gründen nicht zu überzeugen sein, egal wie gut die neuroökonomische Forschung aufgestellt ist. Verhaltensökonomien aber sind aufgrund ihres Paradigmas potenzielle Unterstützer. Stoßen sie jedoch auf die den Kritiken zugrundeliegenden Probleme, werden auch sie keine Neuroökonomien. Mit der Lösung der Probleme können die Neuroökonomien aber tatsächlich weitere Wissenschaftler (darunter potenziell auch die nicht-systematischen Kritiker)

auf ihre Seite bringen und damit das künftige Verhältnis zwischen den Vertretern der Mutterwissenschaften genauso beeinflussen wie die Reputation und Relevanz der Neuroökonomie als Wissenschaft. Dann hat die Neuroökonomie auch eine Chance, sich als Wissenschaft zu etablieren, denn der Themenkomplex kognitive Neurowissenschaften wird nicht einfach wieder verschwinden, dafür sind die Publikationszahlen in diesem Gebiet in den vergangenen zwanzig Jahren zu stark angewachsen, wie eine Recherche beispielsweise auf PubMed Central (PMC), dem Repositorium der U.S. National Institutes of Health's National Library of Medicine (NIH/NLM) ergibt, das einen Anstieg um mehr als das Zwanzigfache verzeichnet.⁴ Ein Grund mehr für Neuroökonom*innen, sich nicht zu schnellen Publikationen und Marketing Bloopers, wie Harrison sie nannte (s. o. Abschn. 6.1.1), hinreißen zu lassen, sondern die Neuroökonomie für die Zukunft auf solide Beine zu stellen.

Open Access Dieses Kapitel wird unter der Creative Commons Namensnennung 4.0 International Lizenz (<http://creativecommons.org/licenses/by/4.0/deed.de>) veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Kapitel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.



⁴ Eine Recherche mit Suchwort „Cognitive Neurosciences“ bringt 126 Publikationen für das Jahr 2000 und 2685 Publikationen für das Jahr 2020 [<https://www.ncbi.nlm.nih.gov/pmc/?term=cognitive+neurosciences>, abgerufen am 08.12.2020].

Literaturverzeichnis

- Abbinck, K./Brandts, J. (2010) Drei Oligopolexperimente. In: Ockenfels, A./Sadrieh, A. (eds.) *The Selten School of Behavioral Economics. A Collection of Essays in Honor of Reinhard Selten*, Heidelberg etc.: Springer
- Albers, W. et al., ed. (1997) *Understanding Strategic Interaction: Essays in Honor of Reinhard Selten*, Berlin etc.: Springer
- Alt, J. E. et al., ed. (1999) *Competition & Cooperation: Conversations with Nobelists about Economics and Political Science [...]*, New York: Russell Sage Foundation
- Baars, B. J./Gage, N. M. (2010) *Cognition, Brain, and Consciousness: Introduction to Cognitive Neuroscience*, Amsterdam: Academic Press
- Bar-Hillel, B. (1973) On the Subjective Probability of Compound Events. *Organizational Behavior and Human Performance*, 9, 396–406
- Bechtel, W. (1988) *Philosophy of Science. An Overview for Cognitive Science*, Hillsdale: Lawrence Erlbaum Associates
- Berg, J. et al. (1995) Trust, Reciprocity, and Social History. *Games and Economic Behavior*, 10 (1), 122–142
- Bernheim, B. D. (2009) The Psychology and Neurobiology of Judgement and Decision Making: What's in it for Economists?. In: Glimcher, P. W. et al., ed. *Neuroeconomics. Decision Making and the Brain*, Amsterdam etc.: Elsevier Academic Press, 115–125
- Bhatt, M./Camerer, C. F. (2005) Self-Referential Thinking and Equilibrium as States of Mind in Games: fMRI Evidence. *Games and Economic Behavior*, 52, 424–459
- Broca, P. (1861) Remarques sur le siège de la faculté du langage articulé, suivie d'une observation d'aphémie (perte de la parole). *Bulletins de la société anatomique*, 6, 330–357
- Büchel, C./Karnath, H.-O./Thier, P. (2012) Methoden der kognitiven Neurowissenschaften. In: Karnath, H.-O./Thier, P., ed. *Kognitive Neurowissenschaften*, Berlin: Springer, 9–32
- Cabeza, R./Nyberg, L. (2000) Imaging Cognition II: An Empirical Review of 275 PET and fMRI Studies. *Journal of Cognitive Neuroscience*, 12 (1), 1–47
- Camerer, C./Loewenstein, G./Prelec, D. (2004) Neuroeconomics: Why Economics Needs Brains. *Scand. J. of Economics*, 106 (3), 555–579
- Camerer, C. (2005a) Neuroeconomics. *Bulletin of the American Academy of Arts and Sciences*, 58 (4), 12–15
- Camerer, C./Loewenstein, G./Prelec, D. (2005b) Neuroeconomics: How Neuroscience Can Inform Economics. *Journal of Economic Literature*, 43, 9–64

- Camerer, C. F. (2005c) *What is Neuroeconomics?* [online]. Abrufbar unter: <http://www.people.hss.caltech.edu/~camerer/web_material/n.html> [abgerufen am 09.08.2017]
- Camerer, C. F. (2007) Neuroeconomics: Using Neuroscience to Make Economic Predictions. *The Economic Journal*, 117 (March), C26–C42
- Camerer, C. (2008a) The Case for Mindful Economics. In: A. Caplin/A. Schotter, ed. *The Foundations of Positive and Normative Economics: A Handbook*, Oxford: Oxford University Press, 43–69
- Camerer, C. (2008b) The Potential of Neuroeconomics. *Economics and Philosophy*, 24 (3), 369–379
- Caplin, A./Schotter, A., ed. (2008) *The Foundations of Positive and Normative Economics: A Handbook*, Oxford: Oxford University Press
- Carter, M./Shieh, J. (2010) *Guide to Research Techniques in Neuroscience*, Amsterdam: Academic Press
- Cartwright, E. (2014) *Behavioral Economics*, London/New York: Routledge
- Chamberlin, E. H. (1948) An Experimental Imperfect Market. *Journal of Political Economy*, 56 (2), 95–108
- Chang, L. J. et al. (2011) Triangulating the Neural, Psychological, and Economic Bases of Guilt Aversion. *Neuron*, 70, 560–572
- Cournot, A. A. (1838) *Recherches sur les principes mathématiques de la théorie des richesses*, Paris: Imprimerie d'Amédée Gratiot
- Deutsche Forschungsgemeinschaft DFG (2015) *Leitlinien zum Umgang mit Forschungsdaten* [online]. Abrufbar unter: <http://www.dfg.de/download/pdf/foerderung/antragstellung/forschungsdaten/richtlinien_forschungsdaten.pdf> [abgerufen am 15.07.2018]
- Dhami, S. (2016) *The Foundations of Behavioral Economic Analysis*. Oxford: Oxford University Press
- European Commission (2019) *H2020 Programme. AGA – Annotated Model Grant Agreement* [online]. Abrufbar unter: <https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/amga/h2020-amga_en.pdf> [abgerufen am 20.07.2020]
- Fehr, E. et al. (1993) Does Fairness Prevent Market Clearing? An Experimental Investigation. *The Quarterly Journal of Economics*, 108 (2), 437–459
- Feyerabend, P. K. (1962) Explanation, Reduction, and Empiricism. In: G. Maxwell/H. Feigl, ed. *Scientific Explanation, Space, and Time*, Minneapolis: University of Minnesota Press, 28–97
- Fließbach, K. (2011) Soziale Präferenzen. In: Reimann, M./Weber, B., ed. *Neuroökonomie. Grundlagen – Methoden – Anwendungen*. Wiesbaden: Gabler, 140–162
- Fodor, J. (1974) Special Sciences (Or: The Disunity of Science as a Working Hypothesis). *Synthese*, 28 (2), 97–115
- Friedman, M. (1953) The Methodology of Positive Economics. In: M. Friedman, ed. *Essays in Positive Economics*, Chicago/London: The University of Chicago Press, 3–43
- Geschwind, N. (1965) Disconnection Syndromes in Animals and Man. *Brain*, 88 (2), 237–294, 88 (2), 585–644
- Geuter, S. et al. (2018) *Effect Size and Power in fMRI Group Analysis*, Preprint: bioRxiv. Persistent Identifier: <https://doi.org/10.1101/295048>
- Glimcher, P. W. et al., ed. (2009a) *Neuroeconomics: Decision Making and the Brain*, Amsterdam etc.: Elsevier Academic Press

- Glimcher, P. W. et al. (2009b) Introduction: A Brief History of Neuroeconomics. In: Glimcher, P. W. et al., ed. *Neuroeconomics: Decision Making and the Brain*, Amsterdam etc.: Elsevier Academic Press, 1–12
- Glimcher, P. W. (2009c) Choice: Towards a Standard Back-pocket Model. In: Glimcher, P. W. et al., ed. *Neuroeconomics: Decision Making and the Brain*, Amsterdam etc.: Elsevier Academic Press, 503–521
- Glimcher, P. W. (2011) *Foundations of Neuroeconomic Analysis*, Oxford etc.: Oxford University Press
- Glimcher, P. W. et al., ed. (2014a) *Neuroeconomics: Decision Making and the Brain: Second Edition*, Amsterdam etc.: Elsevier Academic Press
- Glimcher, P. W. (2014b) Introduction to Neuroscience. In: Glimcher, P. W./Fehr, E., ed. *Neuroeconomics: Decision Making and the Brain: Second Edition*, Amsterdam: Academic Press, 63–75
- Gul, F./Pesendorfer, W. (2008) The Case for Mindless Economics. In: Caplin, A./Schotter, A., ed. *The Foundations of Positive and Normative Economics: A Handbook*, Oxford: Oxford University Press, 3–39
- Güth, W./Strobel, M. (1997) Interview with Elisabeth and Reinhard Selten. In: Albers, W. et al., ed. *Understanding Strategic Interaction: Essays in Honor of Reinhard Selten*, Berlin: Springer, 1–7
- Hansson, S. O./Grüne-Yanoff, T. (2012) Preferences, *Stanford Encyclopedia of Philosophy*, [abgerufen am 31.03.2014]
- Harrison, G. W. (2008a) Neuroeconomics: A Critical Reconsideration. *Economics and Philosophy*, 24 (3), 303–344
- Harrison, G. W. (2008b) Neuroeconomics: A Rejoinder. *Economics and Philosophy*, 24 (3), 533–544
- Harrison, G. W./Ross, D. (2010) The Methodologies of Neuroeconomics. *Journal of Economic Methodology*, 17 (2), 185–196
- Hart, J. (2016) *The Neurobiology of Cognition and Behavior*, Oxford: Oxford University Press
- Hartje, W. (2012) Zur Entwicklung der kognitiven Neurowissenschaften. In: Karnath, H.-O./Thier, P., ed. *Kognitive Neurowissenschaften* [s.u.], 1–7
- Hebb, D. (1949) *The Organization of Behavior: A Neuropsychological Theory*, New York: Wiley
- Head, H. (1926) *Aphasia and Kindred Disorders of Speech*, I–II, Cambridge: Cambridge University Press
- Hermey, G. et al. (2010) *Der Experimentator: Neurowissenschaften*, Heidelberg: Spektrum
- Heukelom, F. (2014) *Behavioral Economics. A History*, Oxford etc.: Oxford University Press
- Hobson, A. (2009) Consciousness and the Brain. In: Laureys, S./Tononi, G., ed. *The Neurology of Consciousness: Cognitive Neuroscience and Neuropathology*, Amsterdam: Academic Press, xi–xii
- Hoyningen–Huene, P. (1989) *Die Wissenschaftsphilosophie Thomas S. Kuhns. Rekonstruktion und Grundlagenprobleme*, Braunschweig/Wiesbaden: Friedr. Vieweg & Sohn
- Huettel, S. A./Song, A. W./McCarthy, G. (2014) *Functional Magnetic Resonance Imaging*, Sunderland: Sinauer
- Ingersoll, J. E. (1987) *Theory of Financial Decision Making*, Lanham: Rowman & Littlefield
- Jevons, W. S. (1871) *The Theory of Political Economy*, London: Macmillan and Co.

- Kahneman, D./Tversky, A. (1973) On the Psychology of Prediction. *Psychological Review* 80, 237–251
- Kahneman, D. (2011) *Thinking, Fast and Slow*. London etc.: Penguin
- Karnath, H.-O./Thier, P., ed. (2012) *Kognitive Neurowissenschaften*, Berlin: Springer
- Kirchgässner, G. (2013) *Homo Oeconomicus. Das ökonomische Modell individuellen Verhaltens und seine Anwendungen in den Wirtschafts- und Sozialwissenschaften*, 4th ed., Tübingen: Mohr Siebeck
- Kuhn, T. (1962) *The Structure of Scientific Revolutions*, Chicago: University of Chicago Press
- LeDoux, J. E. (2012) Evolution of Human Emotion: A View through Fear. In: Hofman, M. A./Falk, D., ed. *Evolution of the Primate Brain: From Neuron to Behavior*, Amsterdam: Elsevier, 431–442
- Lejuez, C. W. et al. (2002) Evaluation of a Behavioral Measure of Risk Taking: The Balloon Analogue Risk Task (BART). *Journal of Experimental Psychology: Applied*, 8 (2), 75–84
- Marshall, A. (1890) *The Principles of Economics*, London: Macmillan
- Meindl, C. (2011) *Methodik für Linguisten. Eine Einführung in Statistik und Versuchsplanung*, Tübingen: Narr
- Moll, J. et al. (2006) Human Fronto-Mesolimbic Networks Guide Decisions about Charitable Donation. *Proceedings of the National Academy of Sciences*, 103 (42), 15623–15628
- Nagel, E. (1961) *The Structure of Science: Problems in the Logic of Scientific Explanation*, London: Routledge & Kegan Paul
- Ockenfels, A./Sadrieh, A. ed. (2010) *The Selten School of Behavioral Economics: A Collection of Essays in Honor of Reinhard Selten*, Heidelberg: Springer
- Ortmann, A. (2008) Prospecting Neuroeconomics. *Economics and Philosophy*, 24 (3), 431–448
- Poldrack, R. A. (2006) Can Cognitive Processes be Inferred from Neuroimaging Data?. *Trends in Cognitive Sciences*, 10 (2), 59–63
- Postle, B. R. (2015) *Essentials of Cognitive Neuroscience*, Chichester: Wiley Blackwell
- Rabin, M. (1998) Psychology and Economics. *Journal of Economic Literature*, 36 (1), 11–46
- Rao, H. et al. (2008) Neural Correlates of Voluntary and Involuntary Risk Taking in the Human Brain: An fMRI Study of the Balloon Analog Risk Task (BART). *NeuroImage*, 42 (2), 902–910
- Reimann, R./Weber, B. ed. (2011) *Neuroökonomie. Grundlagen – Methoden – Anwendungen*, Wiesbaden: Gabler
- Richardson, R. C. (1979) Functionalism and Reductionism. *Philosophy of Science*, 46 (4), 533–558
- Roth, H. L. (2014) We Stand on the Shoulders of Giants: The Golden Era of Behavioral Neurology 1860–1950 and Its Relevance to Cognitive Neuroscience Today. In: Chatterjee, A./Coslett, B., ed. *The Roots of Cognitive Neuroscience: Behavioral Neurology and Neuropsychology*, Oxford: Oxford University Press, 11–52
- Ruff, C. C./Huettel, S. A. (2014) Experimental Methods in Cognitive Neuroscience. In: Glimcher, P. W./Fehr, E., ed. *Neuroeconomics: Decision Making and the Brain: Second Edition*, Amsterdam: Academic Press, 77–108
- Saueremann, H./Selten, R. (1959) Ein Oligopolexperiment. *Zeitschrift für die gesamte Staatswissenschaft / Journal of Institutional and Theoretical Economics*, 115 (3), 427–471
- Schaffner, K. F. (1967) Approaches to Reduction. *Philosophy of Science*, 34 (2), 137–147

- Schaffner, K. F. (1969) The Watson-Crick Model and Reductionism. *The British Journal for the Philosophy of Science*, 20 (4), 325–348
- Sharrock, W./Read, R. (2002) *Kuhn: Philosopher of Scientific Revolution*, Cambridge: Polity Press
- Simon, H. (1955) A Behavioral Model of Rational Choice. *The Quarterly Journal of Economics*, 69 (1), 99–118
- Smith, V. L. (1962) An Experimental Study of Competitive Market Behavior. *Journal of Political Economy*, 70 (2), 111–137
- Sober, E. (1999) The Multiple Realizability Argument against Reductionism. *Philosophy of Science*, 66 (4), 542–564
- Sommer, T. (2010) Neuroimaging: neurobildgebende Verfahren. In: G. Hermeijer et al., *Der Experimentator: Neurowissenschaften*, Heidelberg: Spektrum, 209–263
- Thirion, B. et al. (2007) Analysis of a Large fMRI Cohort: Statistical and Methodological Issues for Group Analyses. *NeuroImage*, 35, 105–120
- Tricomi, E. et al. (2010) Neural Evidence for Inequality-Averse Social Preferences. *Nature*, 463, 1089–1092
- Tversky, A./Kahneman, D. (1974) Judgment under Uncertainty: Heuristics and Biases. *Science*, 185 (4157), 1124–1131
- Tversky, A./Kahneman, D. (1982) Judgments of and by Representativeness. In: D. Kahneman et al., ed. *Judgment under Uncertainty: Heuristics and Biases*, Cambridge etc.: Cambridge University Press, 84–98
- Von Struve, G. (1843) *Die Geschichte der Phrenologie*, Heidelberg: Karl Groos
- Wernicke, C. (1874) *Der aphasische Symptomenkomplex: Eine psychologische Studie auf anatomischer Basis*, Breslau: Max Cohn & Weigert
- Wickens, T. D. (2002) *Elementary Signal Detection Theory*, Oxford etc.: Oxford University Press
- Wilkinson, N. (2008) *An Introduction to Behavioral Economics*, Basingstoke: Palgrave Macmillan
- Zaki, J./Mitchell, J. P. (2011) Equitable Decision Making Is Associated with Neural Markers of Intrinsic Value. *Proceedings of the National Academy of Sciences*, 108 (49), 19761–19766

Webseiten

- Braininitiative.nih.gov <<https://www.braininitiative.nih.gov>>, abgerufen am 20.01.2019
- Humanbrainproject.eu <<https://www.humanbrainproject.eu/en/about/overview/>>, abgerufen am 20.01.2019
- PubMed Central (PMC) <<https://www.ncbi.nlm.nih.gov/pmc/?term=cognitive+neurosciences>>, abgerufen am 08.12.2020