

Experimente in der Soziologie?

Über die systematische Verzerrung von Experimentergebnissen aufgrund strategisch agierender studentischer Teilnehmertypen

Zusammenfassung: In der Soziologie wird vermehrt für eine Verwendung von Labor- und Feldexperimenten plädiert. Bis heute stößt die experimentelle Methode aber auf ein gewisses Unbehagen und gehört auch weiterhin nicht zu den Standardverfahren der empirischen Sozialforschung. Ziel des vorliegenden Beitrags ist es, verzerrende soziale Faktoren experimenteller Laborsettings zu benennen und konstruktiv-kritisch zu besprechen. Dazu wird auf Befunde aus eigenen Verteilungsexperimenten (n=213) in Kombination mit anschließenden problemzentrierten Interviews mit Teilnehmenden (n=39) zurückgegriffen. Ursprünglich wurden die Interviews in das experimentelle Setting integriert, um Aufschluss über die zugrundeliegenden Entscheidungsstrukturen der beteiligten Experimententeilnehmer*innen zu gewinnen. Unter anderem wurden Teilnehmer*innen nach ihren persönlichen Erfahrungen mit Experimenten und ihren Teilnahmemotiven befragt. Die Ergebnisse der qualitativen Befragung weisen auf ein in der Literatur bisher vernachlässigtes Problem hin: So zeigen unsere Interviews neben dem in der Literatur antizipierten strategisch-nutzenmaximierenden Typen noch weitere Teilnehmertypologien, die wir als ‚Interessierte Forscher‘ und ‚Kritische Experten‘ bezeichnet haben. Während für den nutzenmaximierenden Typus ‚Homo Oeconomicus‘ die zentrale Teilnahmemotivation die monetären Anreize darstellt, spielen für die ‚Interessierten Forscher‘ erste Erfahrungen mit Experimenten aus Fachinteresse eine zentrale Rolle. Die ‚Kritischen Experten‘ schließlich nutzen die Experimentteilnahme um ihr eigenes Handeln in spezifischen Situationen kritisch zu hinterfragen und eigene Handlungsentscheidungen zu reflektieren. Da Experimente in der Regel mit Studierenden von Universitäten durchgeführt werden, stellen diese Befunde eine weitreichende Herausforderung für die experimentelle Sozialforschung dar und sollten als solche auch in zukünftige konzeptionelle Überlegungen aufgenommen werden.

Stichwörter: Empirische Sozialforschung; Experimente; Interviews; Selbstselektion; Interne und externe Validität

* Dr. Alexander Lenger, Albert-Ludwigs-Universität Freiburg und Universität Siegen, Zentrum für ökonomische Bildung, Kohlbettstr. 15, 57068 Siegen, E-Mail: lenger@zlb.uni-siegen.de

** Dr. Stephan Wolf, Institut für Umweltsozialforschung, Universität Freiburg, Tennenbacher Str. 4, 79106 Freiburg im Breisgau, E-Mail: stephan.wolf@ere.uni-freiburg.de

Experiments in Sociological Research?

On Systematic Biases of Experimental Findings Due to Strategic Behavior of Student Participants

Abstract: An increasing number of sociologists suggest that laboratory and field experiments should play a larger role in sociological research. Until now, experiments do not form part of the methodological core of empirical social research but are rather met with a certain degree of scepticism. This article discusses social factors which systematically distort experimental laboratory results. Our analysis is based on findings from our own experiment on distributional preferences (n=213) combined with semi-structured interviews with participants (n=39). Initially, the interviews were realized to give insights into the reasoning of participants. Among other issues, we asked subjects about their previous experience with experiments and their motivation behind participating. The results of these qualitative interviews point towards a problem which so far is neglected in research: While the ‘homo oeconomicus’ type of participants – motivated mainly by monetary reward – is anticipated in the literature and also found in our study, our results show that (at least) two additional types of subjects can be identified. First, there is the ‘interested researcher’ who already gained some experience with experiments and whose desire to participate in experimental studies is driven by intellectual curiosity. Second, there is the ‘critical expert’ who uses experiments to critically reflect on her own behavior in specific situations. Given that experiments usually employ student participants recruited at universities, these results are relevant for experimental social research and need to be taken into consideration in future experiment-based research.

Keywords: Empirical Social Research; Experiments; Interviews; Self-selection; Internal and external validity

1 Einleitung

Zur Analyse von wirtschaftlichem Verhalten greifen die modernen Wirtschaftswissenschaften inzwischen flächendeckend auf experimentelle Methoden zurück. Es besteht kein Zweifel, dass die Ökonomik heute zu einem erheblichen Teil eine Experimentalwissenschaft ist (Plott 1991). Vermehrt wird auch in der Soziologie dafür plädiert, dass sich die Soziologie Labor- und Feldexperimenten bedienen sollte, um der notwendigen Methodenvielfalt empirischer Sozialforschung gerecht zu werden (Keuschnigg/Wolbring 2014, 2015b). So diskutiert beispielsweise Andreas Diekmann Ergebnisse aus der experimentellen Wirtschaftsforschung und zeigt, inwieweit diese Befunde an klassische soziologische Erkenntnisse anknüpfen und zur Weiterentwicklung der Sozialtheorie beigetragen haben. „Die Soziologie

wäre gut beraten“, so sein Fazit, „Ergebnisse der experimentellen Ökonomie aufzunehmen und in einen Dialog einzutreten“ (Diekmann 2008: 529).

Entsprechend positiv werden inzwischen in den Sozialwissenschaften die Befunde der experimentellen Sozialforschung bewertet, die mit Hilfe von Experimenten versucht, die soziale Wirklichkeit bewusst stilisiert zu duplizieren, soziale Praktiken im Labor zu reproduzieren und somit der empirischen Untersuchung zugänglich zu machen (vgl. exemplarisch Salganik et al. 2006; Diekmann 2008; Schröder 2011; Winter et al. 2012; Berger/Diekmann 2015; Keuschnigg/Wolbring 2015 b). „Experimente sind quasi ‚Wahrheitsfilter‘, die aus einer Menge von theoretisch inspirierten Hypothesen die korrekten Hypothesen heraus sortieren sollen“ (Diekmann 2008: 530).

Gehörten Experimente in der Psychologie und insbesondere in der Sozialpsychologie schon immer zu einer der zentralen Methoden der Datengewinnung und Theoriebildung, so wurde diese Methodik in der zweiten Hälfte des 20. Jahrhunderts schrittweise in die Mikrosoziologie, die Politikwissenschaft und die Wirtschaftswissenschaften integriert (Zimmermann 2015: 20, vgl. auch Jackson/Cox 2013). Die experimentelle Methodik zählt in den Wirtschaftswissenschaften neben statistischen Analysen und theoretischer Modellierung inzwischen zum Kern ökonomischer Mainstream-Verfahren. Insofern kann es kaum überraschen, dass in den vergangenen Jahren insbesondere die experimentelle Wirtschaftsforschung zunehmend auch in soziologischen Analysen Beachtung findet (vgl. grundlegend Diekmann 2008).

Zugleich wird der Einsatz experimenteller Methoden in den Sozialwissenschaften – ungeachtet aller Befunde und methodischer Verbesserungen – kritisiert, weil sich im Gegensatz zu Naturphänomenen im Bereich sozialer Beziehungen keine Kausalgesetze vorfinden lassen, da sich die soziale Praxis erst aus den Bedeutungen, Intentionen und Institutionen konstituiere. So entzögen sich soziale Phänomene per se der experimentellen Kontrolle und artifizielle Experimente seien in der (Groß-)Gesellschaft ohnehin ethisch kaum zu verantworten (Lorenz/Groß 2010: 12). Aufgrund ihrer strittigen externen Validität stößt die experimentelle Methode bis heute insbesondere in der Soziologie auf Widerstand (Keuschnigg/Wolbring 2015 a: 7) und stellt auch weiterhin ein eher randständiges Verfahren in der empirischen Sozialforschung dar (Schnell et al. 1999: 221).

Der vorliegende Beitrag knüpft an die methodische Untersuchung und kritische Prüfung von Laborexperimenten an (Jackson/Cox 2013; Webster/Sell 2014; Berger 2015; das Sonderheft der *Journal of Economic Behavior & Organization* 73, 2010 zu „Issues in the Methodology of Experimental Economics“). Während diese Auseinandersetzung ihren Blick häufig lediglich auf die experimentell untersuchten Entscheidungssituationen richten (vgl. exemplarisch Ball/Cech 1996; Harrison/List 2004; Levitt/List 2007), verfolgen wir im vorliegenden Beitrag im Anschluss an Roger Berger das Ziel, die methodische Betrachtung des Laborexperiments als Teil

einer sozialen Interaktion und somit als einen genuin sozialen Prozess zu reflektieren (Berger 2015: 54).

Ziel des vorliegenden Beitrags ist es, verzerrende soziale Faktoren experimenteller Laborsettings zu benennen und konstruktiv-kritisch zu besprechen. Dazu wird auf Befunde aus eigenen Verteilungsexperimenten ($n=213$) in Kombination mit anschließenden problemzentrierten Interviews mit Teilnehmenden dieser Experimente ($n=39$) zurückgegriffen. Ursprünglich wurden die qualitativen Interviews im Sinne einer Methodentriangulation und Datenvalidierung in das Forschungssetting integriert, um weiteren Aufschluss über die zugrundeliegenden Entscheidungsstrukturen der beteiligten Experimententeilnehmer*innen zu gewinnen. Im Rahmen dieser Befragung wurden die Teilnehmer*innen unter anderem nach ihren persönlichen Erfahrungen mit Experimenten und ihren Teilnahmemotiven befragt. Die Ergebnisse dieser Befragung zeigten ein neues, bisher in der Literatur vernachlässigtes Problem, das wir im weiteren Verlauf des Beitrags beschreiben wollen: So fanden sich als Resultat der qualitativen Befragung neben dem in der Literatur antizipierten strategisch-nutzenmaximierenden Typen noch weitere Typen von Teilnehmer*innen, die wir als ‚Interessierte Forscher‘ und ‚Kritische Experten‘ bezeichnen. Während für den nutzenmaximierenden Typus ‚Homo Oeconomicus‘ die zentrale Teilnahmemotivation die monetären Anreize darstellen, spielen für die ‚Interessierten Forscher‘ erste Erfahrungen mit Experimenten aus Fachinteresse die zentrale Rolle. Die ‚Kritischen Experten‘ schließlich nutzen die Experimentteilnahme dazu, ihr eigenes Handeln in spezifischen Situationen kritisch zu hinterfragen und eigene Handlungsentscheidungen zu reflektieren. Paradoxerweise resultieren diese Typen zu einem gewissen Grade selbst aus der Attraktivität der experimentellen Methode, welche in verschiedenen Fachdisziplinen Eingang in den Lehr- und Lernalltag von Studierenden gefunden hat. Da Experimente in der Regel mit Studierenden von Universitäten durchgeführt werden, stellen diese Befunde eine weitreichende Herausforderung für die experimentelle Sozialforschung dar und sollten als solche auch in zukünftige konzeptionelle Überlegungen aufgenommen werden. Damit schließt der Beitrag an die Diskussion um Selektivitäts- und Reaktivitätsprobleme der experimentellen Forschung an (Berger 2015).

Der Beitrag gliedert sich wie folgt: Zunächst stellen wir die experimentelle Methodik am Beispiel der experimentellen Wirtschaftsforschung vor (Abschnitt 2.1), präsentieren die zentralen methodologischen Grundlagen (Abschnitt 2.2) und arbeiten die wesentlichen Kritikpunkte heraus (Abschnitt 2.3). In einem weiteren Schritt werden unsere eigenen empirischen Befunde dargestellt. Hierzu wird zunächst das zugrundeliegende Experiment und die ergänzende empirische Befragung erklärt (Abschnitt 3.1). Daran anschließend werden drei Typen von Experimententeilnehmer*innen herausgearbeitet (Abschnitt 3.2) und die systematischen Verzerrungen bei Experimenten mit studentischen Teilnehmer*innen benannt (Abschnitt 3.3). Der Beitrag schließt mit einer Diskussion der wichtigsten Ergebnisse und benennt

abschließend die Konsequenzen sowie einige potenzielle Verbesserungsmöglichkeiten für die experimentelle Sozialforschung (Abschnitt 4).

2 Experimentelle Wirtschaftsforschung

2.1 Ultimatum- und Diktatorspiel

Um die Frage nach einer experimentellen Ausrichtung der Soziologie zu verhandeln, lohnt sich zunächst ein vergleichender Blick in die experimentelle Wirtschaftsforschung (*Experimental Economics*). Ziel der experimentellen Wirtschaftsforschung ist es, die in der traditionellen Ökonomik unterstellte Verhaltensannahme eines rationalen und selbstinteressierten Homo Oeconomicus in *kontrollierten* Entscheidungssituationen zu untersuchen, um das ökonomische Entscheidungsmodell zu einer umfassenden Verhaltensökonomik (*Behavioral Economics*) weiterzuentwickeln. Ein bis heute zentrales Experiment zur Untersuchung ökonomischen Verhaltens und der Grenzen des Homo Oeconomicus-Modells stellt das sogenannte *Ultimatumspiel* (Güth et al. 1982) bzw. seine Abwandlung zum *Diktatorspiel* (Kahneman et al. 1986) dar. In seiner einfachsten und ursprünglichen Form wurde das Ultimatumspiel von Güth et al. (1982) wie folgt durchgeführt: Zwei Probanden erhalten einen gewissen Geldbetrag, z.B. 10 Euro, zur gemeinsamen Aufteilung. Eine der Personen nimmt dabei die Rolle des sogenannten *Proposers* ein. Sie schlägt eine Aufteilung des Betrages auf die beiden Versuchsteilnehmer vor. Die zweite Person entscheidet nun als *Responder* darüber, ob sie die vom Proposer angebotene Verteilung des Gesamtbetrages akzeptiert. Falls ja, erhalten beide den vom Proposer vorgesehenen Anteil; falls nicht, so erhalten beide eine Auszahlung von 0 Euro – der Gesamtbetrag geht also den beiden Spielern ‚verloren‘. Gemäß des neoklassischen Rationalwahlmodells, d.h. des rein eigeninteressierten Homo Oeconomicus, sollte der Proposer eine Aufteilung von 9,99 Euro für sich und 1 Cent für den Responder vorschlagen. Der Responder sollte entsprechend der ökonomischen Standard-Logik das Angebot akzeptieren, da 1 Cent einen höheren Betrag darstellt als 0 Cent, die er oder sie im Falle der Ablehnung erhielte. Dies antizipierend würde der rein eigeninteressierte Proposer den minimal nötigen Betrag anbieten, der die Zustimmung des anderen garantiert – eben 1 Cent. Jedes Angebot höher als 1 Cent wäre irrational, da es den Gewinn des Proposers schmälert und für die Zustimmung des Responders nicht notwendig ist.

Empirisch zeigt sich jedoch, dass meist eine 50:50-Aufteilung vorgeschlagen und von der Gegenseite akzeptiert wird. Ein Vorschlag des Proposers, mehr als 6 Euro für sich zu behalten und weniger als 4 Euro dem Responder zu überlassen, wird mit hoher Wahrscheinlichkeit abgelehnt (vgl. Tisserand 2016). Aus Sicht von Verhaltensökonomien zeigt dieses Experiment, dass Individuen über stabile Fairness-Präferenzen verfügen, die dazu führen können, dass auf individuelle Gewinne verzichtet wird, wenn dies zur Stabilisierung einer ‚fairen Verteilung‘ führt. Um diesen Punkt eingehender zu studieren, variierten Kahneman, Knetsch und Thaler (1986) das

ursprüngliche Setting wie folgt zum Diktatorspiel: Wie gehabt legt der Proposer eine Aufteilung vor, die zweite Person muss jedoch die Aufteilung stets akzeptieren. Der Proposer ist also nicht länger durch das potentielle Veto eines Responders beschränkt. Doch selbst unter dieser Bedingung, die dem Proposer vollständige Entscheidungsmacht einräumt, werden nur äußerst selten Aufteilungen realisiert, bei denen die zweite Person nichts erhält. Zwar steigt die Ungleichheit zugunsten des unbeschränkten Proposers, Aufteilungen jenseits 80:20 bleiben aber selten. Als Erklärungsfaktor für das Abweichen von reinem Eigennutz bleibt aus Sicht der Verhaltensökonomik ein signifikantes Maß an Altruismus. In die Terminologie der Ökonomik übersetzt: Menschen verfügen grundlegend über soziale Präferenzen (*Other-Regarding Preferences*). Unter welchen Umständen diese nicht-egoistischen Neigungen in welchem Umfang zum Tragen kommen, wurde in unzähligen Abwandlungen dieser beiden einfachen Experimente getestet.¹

2.2 Methodologische und konzeptionelle Grundlagen

Hinter solchen sozialwissenschaftlichen Labor- und Feldexperimenten steht die Vorstellung, dass sich unter den Bedingungen eines kontrollierten Experiments einzelne ökonomische Situationen und soziale Interaktionen so darstellen lassen, dass die Wirkungszusammenhänge sowie deren Folgen in Isolation analysiert werden können. Dabei sollen die gewonnenen Ergebnisse und Schlussfolgerungen durch wiederholte Durchführung der Experimente auf ihre Validität und Reliabilität hin überprüfbar sein. Diese Replizierbarkeit stellt aus Sicht der experimentellen Ökonomik den wesentlichen Vorteil ihres Ansatzes dar. In der ungefilterten Realität ist die empirische Überprüfung von Hypothesen nicht möglich, da die wirklichen Rahmenbedingungen einem ständigen Wechsel unterliegen (Schoefer 2005: 28).

Aus dieser Ausgangslage lassen sich zwei wesentliche Ziele experimenteller Methoden ableiten: Zum einen sollen *bestehende Theorien* hinsichtlich ihres Erklärungsgehaltes für das tatsächlich beobachtbare Verhalten überprüft werden. Hierzu werden sogenannte *Hypothesenexperimente* durchgeführt, die eine bestehende Hypothese auf ihre Konsistenz hinsichtlich ihres Erklärungsgehaltes für ökonomische Entscheidungen und soziale Interaktionen prüfen. Zum anderen ist die Konstruktion *neuer Theorien* eine zentrale Aufgabe der experimentellen Sozialforschung. Hierzu werden sogenannte *Erkundungsexperimente* durchgeführt, d.h. es wird aufgrund der Ver-

1 Etwa zum Einfluss von Geschlecht (Saad/Gill 2001; Solnick 2001; Camerer 2003; García-Gallego et al. 2012), kultureller Unterschiede und ethnischer Zugehörigkeit (Henrich 2000; Camerer 2003; Oosterbeek et al. 2004; Henrich et al. 2005; Chuah et al. 2007), Alter (Camerer 2003; Hoffmann/Tee 2006; Harbaugh et al. 2007; Bailey et al. 2013), dem Effekt unterschiedlicher Informationssettings und Kommunikationsmöglichkeiten (Croson 1996; Güth/van Damme 1998; Camerer 2003), der Rolle von Gerechtigkeitsvorstellungen und sozialen Normen (Güth/van Damme 1998; Nowak et al. 2000; Konow 2003), Mehrpersoneninteraktionen (Bahr/Requate 2013; Wolf/Dron 2015) etc. Für eine Übersicht über die Verwendung des Ultimatumspiels in den vergangenen dreieinhalb Jahrzehnten seit dessen Einführung durch Güth et al. (1982), siehe z.B. Güth/Kocher (2014) und Tisserand (2016).

suchsergebnisse eine ‚neue‘ Theorie für das beobachtbare Verhalten der Wirtschaftssubjekte generiert (Smith 2008; Croson/Gächter 2010).

Die experimentelle Methode bietet eine Möglichkeit, kausale Beziehungen zu untersuchen. Dabei wird postuliert, dass zwischen zwei Parametern ein Ursache-Wirkungs-Verhältnis besteht: Die Änderung des einen Parameters (X) führt notwendig zu einer Veränderung des anderen (Y). Eine einfache Beobachtung der empirischen Realität ist dabei nicht ausreichend, da aufgrund sich permanent ändernder Rahmenbedingungen keine eindeutigen Aussagen getroffen werden können. Daher isoliert die empirisch arbeitende Forscherin die zu untersuchende Beziehung in einem Experiment, bei welchem alle anderen möglichen Ursachen einer Veränderung von Y weitestgehend ausgeschlossen werden. Dies geschieht dadurch, dass alle Einflussparameter außer X – soweit umsetzbar – konstant gehalten werden (Ceteris-Paribus-Bedingung), während X systematisch variiert wird. Etwaige Veränderungen von Y lassen sich dann eindeutig auf die Variation von X zurückführen und entsprechend qualitativ und/oder quantitativ beschreiben. Die kausale Wirkungskette ‚Wird X um den Wert x geändert, dann ändert sich Y um den Wert y ‘ gilt dann als etabliert.²

Das Labor bietet gemäß dieser Logik aus Sicht der experimentellen Sozialforschung die notwendigen Bedingungen, um in einem Experiment alle Variablen außer der unabhängigen Variable konstant zu halten, da der Einfluss störender Faktoren entweder verhindert oder zumindest kontrolliert bzw. gemessen werden kann.³ Treffend fasst diese Position Diekmann zusammen: „Das experimentelle Design ist nach wie vor der ‚Königsweg‘ zur Untersuchung kausaler Effekte und zur Prüfung von Kausalhypothesen“ (Diekmann 2015: 3; Keuschnigg/Wolbring 2015 a: 7). Hierbei ist eine große Spannbreite an experimentellen Designs zu konstatieren, da je nach Fragestellung unterschiedliche Techniken und Richtlinien angebracht sind, nach denen ein Experiment vorbereitet, durchgeführt und ausgewertet werden sollte. Einen Überblick zu möglichen Designelementen und Verwendungsfällen bieten

- 2 Dabei müssen die Parameterwerte nicht notwendigerweise in kontinuierlichen, reellen Zahlen gemessen werden. Oftmals ist eine qualitative Beziehung ‚Das Vorhandensein von X beeinflusst Y‘ ausreichend bzw. die Feststellung der Richtung eines Effekts (eine Erhöhung von X steigert/senkt den Wert von Y); eine noch exaktere Aussage ist in solchen Fällen meist nicht möglich. Zwischen diesen beiden Polen der ‚Quantifizierung‘ stehen Fälle, in denen sich die Wirkungskette immerhin ordinal messen lässt, also in diskreten Schritten (z.B. in einer Skala von ‚1 = gering‘ bis ‚5 stark‘ sowohl bzgl. Einwirkung von X als auch hinsichtlich Reaktion von Y).
- 3 Ökonometrische Methoden, die ebenfalls der Untersuchung kausaler Zusammenhänge dienen, sind auch vom Problem betroffen, dass sich aufgrund gleichzeitiger Veränderung mehrerer Variablen oft kein eindeutiges Bild ergibt. Entsprechend müssen technisch aufwendige Verfahren angewandt werden, um Daten der nicht-experimentellen Welt sinnvoll untersuchen zu können. Die experimentelle Forscherin versucht dagegen nicht *ex post* eine ‚Entwirrung‘ der Daten, sondern möchte bereits *ex ante* wesentliche Quellen multikausaler Einflüsse durch ein geeignetes experimentelles Design ausschließen.

Davis/Holt (1993), Friedman/Sunder (1994), Shadish et al. (2002) oder Friedman/Cassar (2004).

Ein konstituierendes Merkmal ökonomischer Experimente ist u.a. die monetäre Entlohnung der Teilnehmer*innen (Falk 2001: 301). Anders als klassische Umfragestudien oder sozialpsychologische Experimente haben die im ökonomischen Experiment getroffenen anreizbasierten Entscheidungen somit faktische monetäre Konsequenzen. Indem damit auf die tatsächlich beobachtbaren Präferenzen (*revealed preferences*) abgestellt wird, wird die Realität besser abgebildet als in alternativen Untersuchungsdesigns. Zudem werden ökonomische Experimente in der Regel ohne Täuschung durchgeführt, d. h. ohne bewusste Falschinformation der Teilnehmenden (für Ausnahmen siehe Weimann 1994). Konkret bedeutet dies, dass die angekündigten Beträge auch tatsächlich ausgezahlt werden und dass das Experiment sich auch tatsächlich auf den angekündigten Sachverhalt bezieht. Im Gegensatz dazu zielen in der Sozialpsychologie Experimente vielfach darauf ab, die eigentliche Fragestellung zu verschleiern, um Aussagen über die wirklichen Einstellungen der Versuchsteilnehmer*innen in sozial unerwünschten oder problembehafteten Situationen treffen zu können (Aronson et al. 2004). So wurde beispielsweise im Milgram-Experiment die Bereitschaft durchschnittlicher Personen getestet, autoritären Anweisungen auch dann Folge zu leisten, wenn sie in direktem Widerspruch zu ihren moralischen Intentionen stehen. Das Setting bestand darin, dass eine Versuchsperson einer anderen Versuchsperson (einem Schauspieler) bei Fehlern in der Zusammensetzung von Wortpaaren jeweils einen elektrischen Schlag versetzte. Ein Versuchsleiter (ebenso ein Schauspieler) gab dazu Anweisungen, wobei die Intensität des elektrischen Schlages fortlaufend gesteigert wurde. In Wirklichkeit wurden jedoch keine Stromschläge gegeben (Milgram 1974; Burger 2009).

2.3 Kritik an sozialwissenschaftlichen Experimenten

Die Kritik an der experimentellen Wirtschaftsforschung, die sich beispielsweise auf Aufbau und Ergebnisse des Ultimatum-Spiels bezieht, betrifft vor allem methodologische Fragen (Fehr/Schmidt 1999). Dabei wird die Gewährleistung der klassischen Gütekriterien einer jeden statistisch fundierten Forschung in Frage gestellt: Generalisierbarkeit, Objektivität, und Validität des Forschungssettings und der Forschungsergebnisse.

Jedes experimentelle Design sieht sich stets mit der Polarität „analytische Isolation vs. realistische Kontextualisierung“ konfrontiert (Held et al. 2003: 10). Laborexperimente sind bewusst so vereinfachend gestaltet, dass sich Ergebnisse möglichst eindeutig auf kausale Ursachen zurückführen lassen und (unkontrollierbare) Verzerrungen weitestgehend vermieden werden. Ziel ist hier, größtmögliche interne Validität zu gewährleisten. Die Isolierung einzelner (kausaler) Faktoren erfordert aber entsprechend, andere, in der Realität ebenfalls vorhandene, Einflüsse auszuschließen oder aufwändig zu kontrollieren. Dieser Schritt der bewussten Herstellung einer

artifiziellen Laborumgebung geht notwendigerweise einher mit dem Verlust externer Validität: Ob die Forschungsergebnisse außerhalb des isolierten Laborsettings und damit eine allgemeine Gültigkeit besitzen, bleibt zunächst offen (Shadish et al. 2002: 18; vgl. zum Problem der externen Validität auch Henrich et al. 2010).

Im Kern besagt die Kritik an der experimentellen Methode, dass ein Experiment niemals verallgemeinerbare Aussagen produzieren kann, eben weil es sich um ein Experiment in artifizieller Umgebung handelt. Am eindrücklichsten fassen Schnell et al. diese Position gegenüber ‚technischen Problemen‘ zusammen:

„Will man experimentelle Resultate auf größere Personengruppen (Grundgesamtheiten) generalisieren, so sind spezielle Auswahlverfahren (insbesondere Zufallsauswahl) zur Rekrutierung der Versuchspersonen unumgänglich. Sollen die Ergebnisse eines Experiments z.B. für die gesamte Bevölkerung der Bundesrepublik verallgemeinert werden, so ist es ratsam, eine entsprechende Auswahl (etwa eine Stichprobe mit vielleicht 1500 Personen) dem Experiment zugrunde zu legen. Dies wäre mit großem Aufwand verbunden, da entweder die ausgewählten Personen in ein Labor kommen müssten oder aber die oft umfangreichen Hilfsmittel ins Feld zu transportieren wären.“ (Schnell et al. 1999: 221)⁴

Croson und Gächter (2010) weisen ebenfalls treffend auf diesen Trade-off zwischen Allgemeinheit und Präzision hin: Sehr präzise Experimente liefern präzise Vorhersagen, aber dazu müssen viele Parameter hinsichtlich ihrer Größen klar bestimmt sein. Nur ein solches Setting erlaubt recht exakte Vorhersagen. Im Gegenzug heißt das aber auch, dass die Vorhersage nur in einem sehr engen Parameterrahmen gilt. Sehr allgemeine Aussagen beziehen sich hingegen auf einen großen Intervallbereich von Parametern oder unterschiedlichen persönlichen Qualitäten, was zu Lasten der Exaktheit der Vorhersage gehen muss. Die Kritik an den ‚unpräzisen‘ Experimenten ist insofern etwas verfehlt: Da Experimente darauf abzielen, sehr allgemeine Aussa-

- 4 In der Tat ist in der experimentellen Ökonomik eine Verlagerung der Aktivitäten hin zu Feldexperimenten zu beobachten (vgl. Carpenter et al. 2005; Levitt/List 2009). Da in den vergangenen Dekaden sehr viele grundlegende, einfache Settings im Labor getestet wurden und die Variationsmöglichkeiten oftmals ausgereizt scheinen (vgl. die Unzahl an Abwandlungen des Ultimatum- und Diktatorspiels), erhoffen sich viele Ökonom*innen durch die kontrollierte Erweiterung des Kontextes hin zu Feldstudien neue Erkenntnisse (Harrison/List 2004; Levitt/List 2009). Damit sollen die beschriebenen Lücken bezüglich externer Validität schrittweise geschlossen werden. Allerdings zeigt sich beim Übergang von Labor- zu Feldexperimenten eben genau das Problem, das zunächst zur Verbreitung der Labormethode geführt hat: Die Bedingungen im Feld sind oft so komplex, dass die Aufrechterhaltung des eigentlichen Settings, das am Laboraufbau orientiert ist, oftmals nicht gelingt (Roe/Just 2009). Um nur ein paar der auftretenden Probleme zu nennen: Die wesentlich mühsamere Instruktion nicht-studentischer Teilnehmer*innen führt zu unerwünschten Lerneffekten, die Anpassung von Auszahlungshöhen an Gruppen mit großer Einkommensheterogenität unterminiert die Vergleichbarkeit der Anreize, und die im Feld weit stärker sich ständig wandelnden Umgebungsparameter verzerren die ‚Messergebnisse‘ ganz grundsätzlich auf unkontrollierte Weise. Insgesamt zeigt sich, dass die Erhöhung der externen Validität bei Feldexperimenten oft nur auf Kosten der internen Validität zu erreichen ist, was dem grundlegenden Ansatz der isolierten Betrachtung von Phänomenen unter ‚ansonsten gleichen Bedingungen‘ zuwiderläuft. Insofern kann davon ausgegangen werden, dass die experimentelle Methode per se nur begrenzt in der Lage ist, ein beliebig großes Maß an externer Validität zu gewährleisten.

gen zu treffen, muss eine solche Konstruktion notwendig zu Lasten der Detailvorhersagekraft gehen. In einem zweiten Schritt aber werden die Parameter immer enger eingegrenzt, um von recht unpräzisen, aber allgemeingültigen Aussagen zu immer exakteren, aber spezifischeren Aussagen zu kommen.

Aber auch die interne Validität, also ob in der experimentellen Situation das gemessen wird, was gemessen werden soll, wird häufig angezweifelt. „Nobody thinks that laboratory results on the Dictator Game will predict how much people will donate from their wallets to strangers they pass in the street. The range of application of results in the Ultimatum Game is similarly limited.” (Binmore/Shaked 2010: 91) Binmore und Shaked beispielsweise argumentieren, dass bereits kleine Änderungen der Versuchsanordnung im Ultimatumspiel zu unterschiedlichen Ergebnissen führen können. Ganz allgemeinen sprächen nach Ansicht der beiden Autoren kulturelle Unterschiede gegen eine erfolgreiche Verallgemeinerung experimenteller Ergebnisse. Gerade im Bereich der Herausbildung individueller Fairnesspräferenzen weisen Menschen je nach Sozialisation ganz unterschiedliche Vorstellungen auf (vgl. hierzu auch Henrich et al. 2001). Solche sozialpsychologischen Mechanismen gilt es nach Binmore und Shaked stets zu berücksichtigen, was spezifische experimentelle Forschungsergebnisse per se relativiert: „[W]hat subjects count as fair depends on a whole range of contextual parameters, including perceived need and prior investment of effort. A context-free theory like inequity aversion cannot therefore apply to all situations in which fairness attitudes matter” (Binmore/Shaked 2010: 91).

Zur fehlenden Repräsentativität von psychologischen und wirtschaftswissenschaftlichen Experimenten aufgrund der Tatsache, dass die Mehrzahl aller Experimentteilnehmer*innen selbst Studierende dieser Fächer sind, liegen inzwischen zahlreiche Befunde vor (Arnett 2008; Henrich et al. 2010; Cleave et al. 2013; Falk et al. 2013; Slonim et al. 2013). Insbesondere wird kritisiert, dass die Probandengruppen nicht nach dem Zufallsprinzip zusammengestellt werden (Cook/Campbell 1979; Shadish et al. 2002), sondern nach dem Prinzip der Verfügbarkeit (einem sogenannten *Convenience Sample*). So wird in vielen Fällen in der Universität rekrutiert oder die Teilnehmer*innen an eigenen Lehrveranstaltungen (zwangs-)verpflichtet. Exemplarisch sei angeführt, dass etwa 96 Prozent aller Testpersonen aus westlichen Industriestaaten stammten und 67 Prozent der amerikanischen Probanden von psychologischen Studien selbst wiederum Psychologiestudent*innen waren (Henrich et al. 2010). Die mangelnde externe Verallgemeinerbarkeit dieser Ergebnisse liegt auf der Hand (Shadish et al. 2002; Slonim et al. 2013). Begegnet wird dieser Problematik, indem der Fokus auf die Maximierung der internen Validität gelegt wird und dazu Probanden „standardmäßig zufällig aus einem Pool prinzipiell interessierter Personen“ (Berger 2015: 56) gezogen werden.

Roger Berger hat die zentralen Aspekte der Probandenrekrutierung zusammengefasst. So erfolgt die Anmeldung zu diesem Pool durch *Selbstselektion* (Aushänge, mündliche Ankündigungen etc.). Meistens handelt es sich um Studierende der Ein-

richtung, in der die Experimente durchgeführt werden. „Damit liegt normalerweise eine relativ homogene Population bezüglich Alter und Bildung vor, selbst wenn darauf geachtet wird, Bewerber aus möglichst verschiedenen Fächern zu rekrutieren.“ (Berger 2015: 56) In der Literatur wird zudem davon ausgegangen, dass sich die Proband*innen nicht strategisch zu den Experimenten verhalten:

„Durch die zufällige Auswahl aus dem Bewerberpool wird aber zumindest vermieden, dass sich Probanden speziell für Experimente melden, die z.B. besonders hohe Gewinne versprechen oder besonders interessant erscheinen. In beiden Fällen würde die Stichprobe zusätzlich durch Selbstselektion verzerrt: entweder durch übermäßig monetär motivierte oder durch inhaltlich besonders interessierte Bewerber. Angestrebt wird jedoch eine Population, die den wissenschaftlichen Aspekten des Experiments neutral gegenübersteht und einzig ein monetäres Interesse an den Auszahlungen hat.“ (Berger 2015: 56)

Diese Argumentation wird im weiteren Verlauf empirisch widerlegt und entsprechend in der Diskussion noch einmal kritisch aufgegriffen.

3 Empirische Befunde

Nachfolgend werden die Ergebnisse der rekonstruktiven Befragung zu den Teilnahmemotiven von Experimententeilnehmer*innen vorgestellt. In einem ersten Schritt wird kurz das methodische Vorgehen skizziert (Abschnitt 3.1). Daran anschließend werden die zentralen Befunde dargestellt. Es können drei Typen der Teilnahmemotivation unterschieden werden (Abschnitt 3.2), die anschließend mit verschiedenen Ansätzen zur Erklärung von systematischen Verzerrungen abgeglichen werden (Abschnitt 3.3).

3.1 Methodik der empirischen Befragung

Die Frage nach den Teilnahmemotiven und Teilnahmetypologien wurde als Ergebnis einer *explorativen* Annäherung auf der Basis problemzentrierter Interviews mit Experimententeilnehmer*innen erhoben und inhaltsanalytisch ausgewertet. Die qualitative Befragung wurde wie folgt realisiert: Vor Beginn eines ökonomischen Verteilungsexperiments⁵ (eine Kurzzusammenfassung zum Aufbau und Ablauf des Expe-

5 Die jeweils neun Teilnehmer*innen einer Gruppe hatten aus einer Reihe von Verteilungsschemata per Mehrheitsentscheid eines auszuwählen. Dabei unterschieden sich die Schemata bzgl. der Einkommensspreizung der neun zu besetzenden Positionen von vollkommen gleichverteilt bis stark ungleich. Nach der kollektiven Entscheidung für ein Schema nahmen alle Proband*innen an einem Produktionsspiel mit individuell zu bewältigenden Aufgaben teil. Die im Produktionsschritt erzielten Punkte entschieden dann über die Position, die ein Proband im Verteilungsschema einnahm; in Verbindung mit dem gewählten Schema bestimmte sich so die individuelle Auszahlung. Alle Teilnehmer*innen wurden persönlich über ihren Erfolg informiert. In einer zweiten Runde konnten die Proband*innen nochmals über das Einkommensschema abstimmen, welche die Auszahlungen für die zweite Produktionsrunde bestimmt. Der Produktionsschritt erfolgte analog. In der Kontrollanordnung wurden die Auszahlungen beider Runden einfach addiert; in der Treatment-Anordnung wurde der in Runde 1 erzielte Betrag mit 90% gewichtet, der aus Runde 2 mit 10%. Die Ergebnisse von insgesamt 23 Gruppen zeigen, dass die Teilnehmer*innen im Treatment-Design meist Gleichverteilung in beiden Run-

rimentes findet sich in Fußnote 5, für Details siehe Wolf/Lenger 2014) wurden die sozioökonomischen Faktoren der Versuchsteilnehmer*innen mittels eines Strukturfragebogens erhoben und gefragt, ob sie noch für ein kurzes qualitatives Interview nach Spielende zur Verfügung stehen würden. Am Experiment nahmen insgesamt 213 Student*innen teil, es wurden 39 problemzentrierte Interviews geführt (Leitfaden im Anhang).

Die Samplebildung folgte dem Prinzip der *maximalen strukturellen Variation* (Kleinling 1982). Ein grundlegendes Merkmal qualitativer Samples ist die Kontrastierung relevanter Fälle aus dem empirischen Feld (bewusstes kontrastierendes bzw. komparatives Sampling). Hierüber wird erreicht, ein ‚breites Spektrum‘ an Relevanzsystemen aufzunehmen und somit der Heterogenität des Analysefeldes gerecht zu werden (Kelle/Kluge 1999). Über die Kontrastierung des Feldes wird es möglich, in den Daten zugrundeliegende Muster von Aussagensystemen zu rekonstruieren. Das Ziel rekonstruktiver Studien ist keine statistische Repräsentativität, sondern die phänomenologische, d.h. umfassende und vielschichtige, Repräsentation komplexer sozialer Wirklichkeiten, die hermeneutisch rekonstruiert und als Deutungsmuster dargestellt werden (Helfferich 2011; Kruse 2014).

Konkret auf das Forschungsprojekt bezogen bedeutet dies, dass aus der Gruppe der zur Verfügung stehenden Interviewpersonen Akteure mit unterschiedlichen Positionen und Positionierungen für die Samplebildung herangezogen wurden. Dazu zählen das Abschneiden im Experiment (hohe oder niedrige Auszahlung im Spiel), das experimentelle Verhalten (Verteilungspräferenzen und Reformneigungen), das Geschlecht, der Studiengang sowie die soziale Herkunft. Ausgewertet wurden sämtliche Interviews, sowohl in einfacher inhaltsanalytischer (Mayring 2015) als auch in rekonstruktiver Weise (Kruse 2014). Für die vorliegende Publikation wurde die gesprochene Sprache zur besseren Lesbarkeit ‚geglättet‘, es wurden also sprachliche Fehler und für die gesprochene Sprache typische Füllwörter am Satzanfang gestrichen (Kuckartz et al. 2008: 27-28). Die Satzform hingegen wurde auch mit syntaktischen Fehlern beibehalten. Die Zeichensetzung erfolgte unter Berücksichtigung der Sprachmelodie nach den Regeln der deutschen Rechtschreibung.

3.2 Drei Typen der individuellen Teilnahmemotivation

In den Interviews zeigen sich deutlich unterschiedliche Typen hinsichtlich der Teilnahmemotivation am Experiment. Bei der Typenbildung folgen wir dem Verfahren von Kelle und Kluge (1999). Ausgewertet wurden die Antworten zum zweiten Fragenblock des Leitfadens zur Bewertung von ökonomischen Experimenten. Auch wenn die Motive sich zum Teil überschneiden, so konnten die Antwortstrukturen drei Subkategorien dominanter Motivlagen zugeordnet werden (vgl. Tabelle 1). In den Fällen, in denen mehrere Motive zur Experimentteilnahme ausgeführt wurden

den wählten, während im Control-Design oft zunächst Gleichverteilung und in der anschließenden Runde größtmögliche Ungleichheit gewählt wurde.

(Geld und Interesse), wurde die Person dem im Interview dominierenden Motiv zugeordnet. In der Regel konnte dies über das Vorliegen einer Gewinnmaximierung zugeordnet werden. Eine Motivmischung kam aber nur in 20 Prozent aller Interviews überhaupt zum Tragen.

Tabelle 1: Teilnahmetypen im Experiment

	Typ 1: Homo Oeconomicus	Typ 2: Interessierte Forscher	Typ 3: Kritische Experten
Teilnahmemotiv	Gewinn	Fachinteresse	Handlungsinteresse
Dominierendes Narrativ	Monetärer Gewinnorientierung	Professionelles Erkenntnisinteresse	Persönliche Werthaltung
Gewinnorientierung	Gewinnmaximierung	--	Gleichverteilung
Anzahl der Fälle	22	12	5
Relative Verteilung	56,4 %	30,0 %	12,8 %

Wir haben diese drei Typen als ‚Homo Oeconomicus‘, als ‚Interessierte Forscher‘ sowie als ‚Kritische Experten‘ bezeichnet. Eine vergleichbare Typologie konnte zuvor von Slonim et al. (2013) beobachtet werden. Prozentual lassen sich – bei angemessener Berücksichtigung der schwierigen Verallgemeinerung dieser sehr kleinen Interviewgruppe – in unserem Sample circa 56 Prozent dem Typ ‚Homo Oeconomicus‘, 30 Prozent dem Typ ‚Interessierte Forscher‘ und 13 Prozent dem Typus ‚Kritische Experten‘ zuordnen. Ausschlaggebend für die unterschiedlichen Motivstrukturen sind der generelle Umgang und die Vertrautheit mit Experimenten.

3.2.1 Typ 1: ‚Homo Oeconomicus‘

Für die Interviewpartner*innen, die wir dem Typus des ‚Homo Oeconomicus‘ zugeordnet haben, stellt ausschließlich Geld einen Anreiz zur Teilnahme an dem Experiment dar. Es kann „EIN BISSCHEN GELD“ verdient werden, wobei stets betont wird, dass es sich um einen „GUTEN STUNDENLOHN“ handelt. Es lohnt sich „VORBEIZUSCHAUEN“ und das Geld „MITZUNEHMEN“. Die Ausrichtung am Geld wird explizit und dezidiert hervorgehoben:

„(.) ALSO ES (.) LAG ABER AUCH HAUPTSÄCHLICH DARAN, DASS MAN VORHER JA NICHT WIRKLICH WUSSTE, UM WAS ES GEHT ALSO ICH (.) KONNTE MIR ÜBERHAUPT NICHTS VORSTELLEN, ICH (.) HATTE ÄNGST, DASS WIR IRGENDWELCHE WIRTSCHAFTLICHEN DISKUSSIONEN FÜHREN MÜSSEN UND ICH TOTAL ABKACKE, WEIL ICH KEINE ÄHNUNG VON WIRTSCHAFT HABE (.) ÄHM DESWEGEN WAR ES EINFACH NUR DAS GELD UND ICH HATTE IRGENDWIE NUR DIE HOFFNUNG, DASS ICH GELD RAUSBESKOMME ((LACHT))“

Im Narrativ dieses Typs erscheint die „ARBEIT“ bzw. „LEISTUNG“ im Experiment weniger anstrengend als alternative Erwerbsmöglichkeiten. Sie orientieren sich am Durchschnittseinkommen und gehen davon aus, dass die meisten anderen Versuchsteilnehmer*innen ebenfalls wegen der monetären Anreize an dem Experiment teilnehmen.

Interessant ist, dass die Angehörigen des Typs ‚Homo Oeconomicus‘ eine hohe Expertise im Umgang mit Experimenten aufweisen und regelmäßig an Experimenten in verschiedenen Fachbereichen, wie beispielsweise der Psychologie, teilnehmen. Häufig werden diese psychologischen Experimente bevorzugt, was „AUCH NAHELIEGEND IST, WEIL EINFACH VIEL MEHR ANGEBOTEN WERDEN“. Eine Präferenz für bestimmte Experimente konnte bei diesem Typus mittels Rückfrage aber nicht identifiziert werden. Diese pragmatische Haltung gegenüber Experimenten führt in der Praxis so weit, dass die Teilnahme an Experimenten zu einem relevanten Faktor für die Lebenshaltungskosten wird und signifikant zum Lebensunterhalt beiträgt:

„ICH HABE SCHON FRÜHER SEHR VIELE EXPERIMENTE GEMACHT, IMMER, ALSO AM PSYCHOLOGISCHEN INSTITUT, ICH HAB VORHER IN [STADT] STUDIERT UND DA GAB ES HALT AUCH EIN PSYCHOLOGISCHES INSTITUT UND (.) ALSO ICH HAB DAMIT HALT TEILWEISE MEIN STUDIUM AUCH FINANZIERT“

Im Gegensatz zu anderen Typen stellt das Experiment für die Personen, die eine ‚Homo Oeconomicus‘-Erzählung aufweisen, kein „SPIEL“ dar, sondern eine ernste Angelegenheit, bei der reales Geld ausbezahlt wird. Dementsprechend agieren die ‚Homo Oeconomicus‘-Proband*innen im Experiment auch äußerst strategisch und ihr Agieren ist auf eine potenzielle Maximierung der Auszahlungssumme ausgerichtet. Um ihre Interessen durchzusetzen nehmen sie notfalls auch die „UNSYMPATHISCHE ROLLE“ eines „KAPITALISTEN“ ein, der eine ausgeprägte Affinität für „WETTBEWERB“ und „KONKURRENZ“ zur Schau trägt.

3.2.2 Typ 2: ‚Interessierte Forscher‘

Das Narrativ, mit dem wir den zweiten Typ beschreiben, ist das der am Thema interessierten Forscherin. Zwar spielt auch hier die finanzielle Auszahlung eine Rolle, stellt aber keineswegs den zentralen Anreiz zur Teilnahme am Experiment dar. Vielmehr besteht das Interesse „MAL SO EIN EXPERIMENT SELBER MITZUMACHEN“ und „ZU SEHEN, WIE ES VERLÄUFT UND VIELLEICHT AUCH IRGENDWANN MAL DAS ERGEBNIS VON DEM EXPERIMENT ZU SEHEN“. Die alleinige Ausrichtung am monetären Interesse fehlt den Interviewten diesen Typs; vielmehr betonen sie das inhaltliche Interesse an Experimenten im Allgemeinen. Proband*innen aus den Wirtschaftswissenschaften oder der Psychologie bringen dabei oft ein ganz fachspezifisches Interesse an Verhaltensexperimenten mit, da sie in ihrem Studium und im Austausch mit ihren Kommiliton*innen ein gewisses Grundwissen über Experimente erworben haben. Entsprechend wollten sie aus fachlicher Neugier „SCHON LÄNGER MAL AN SO EINEM EXPERIMENT TEILNEHMEN“, oder weil beispielsweise

„EINE FREUNDIN JETZT GERADE IN PSYCHO SO EIN EXPERIMENT MACHT“ und das „SCHON SEHR SPANNEND“ ist.

In der Folge finden sich unter den „interessierten Forscher*innen“ eine große Anzahl von Studierenden der Wirtschaftswissenschaften. Ein erheblicher Teil der ökonomischen Fachliteratur befasst sich heutzutage mit verhaltenswissenschaftlichen Ansätzen und Ergebnissen, und durch die Teilnahme an einem Experiment erwarten sich die ‚interessierten Forscher*innen‘ praktische Einblicke in Methodik und Verfahrensweise der experimentellen Ökonomik. Der Anspruch ist also „AUCH MAL SO EIN BISSCHEN HINTER DIE KULISSEN ZU GUCKEN“ und dabei auf bestehendes „VORWISSEN“ rekurrieren:

„ALSO ER [DER DOZENT, DIE VERFASSER] HAT JA AUCH VIEL SO VERHALTENSWISSENSCHAFTLICHE SACHEN IN DEN VORLESUNGEN ERZÄHLT UND SO UND DA HAT NATÜRLICH SO EIN – ALSO ICH WOLLTE HALT MAL SEHEN WIE ES SO EIGENTLICH AUSSIEHT ALSO WIE AUCH SO GEFORSCHT WIRD AN DER UNI [...] DAS HAT MICH SO MOTIVIERT EIGENTLICH“

Dieses Interesse geht teilweise so weit, dass die Experimentteilnahme damit begründet wird, sich selber in Richtung experimenteller Wirtschaftsforschung orientieren zu wollen, also im Sinne einer zukünftigen Studienplanung zu schauen „OB MAN GROB IN DIESE RICHTUNG GEHEN WILL“ und entsprechend „GUCKEN OB DAS FÜR MICH EVENTUELL AUCH MAL IN FRAGE KOMMT“.

Auch diese Gruppe verfügt teilweise über große Vorerfahrung mit Experimenten. Zum Teil haben sich Experimentteilnehmer*innen bereits in der Schule mit solchen „PLANSPIELEN“ auseinandergesetzt und finden Experimente „SUPERINTERESSANT“ und freuen sich die Gelegenheit zu bekommen, mal wieder an einem realen Experiment teilzunehmen:

„JA ICH FIND DAS SUPERINTERESSANT DESWEGEN HAB ICH MICH AUCH FAST GEFREUT, DASS ICH DEN PLATZ [IN DER EXPERIMENT-SESSION, DIE VERFASSER] DANN AUF DIESEM WEG DOCH NOCH BEKOMMEN HABE ALS ERSATZ. ALSO ICH HAB AUCH AUF DEN ANMELDEBOGEN GESCHRIEBEN, DASS ICH FÜR EIN [...] INTERVIEW NOCH BEREIT BIN, AUCH WEIL MICH SOZUSAGEN INTERESSIERT, WAS LETZTENDLICH BEI SO WAS RAUSKOMMT. ALSO FIND ICH SUPER, AUCH PLANSPIELE UND SO WEITER“

Entsprechend ihrer motivationalen Ausrichtung betonen die ‚Interessierten Forscher‘ auch den Spielcharakter des Experiments und weisen im Gegensatz zum ‚Homo Oeconomicus‘ kaum strategisches Verhalten im experimentellen Setting auf: Sie wollen um den Auszahlungsbetrag „ZOCKEN“ und betonen die „EXTERNE GELDQUELLE“. Gemeint ist die Möglichkeit, einen zusätzlichen Beitrag außerhalb ihrer regulären Einkommensverhältnisse (Eltern, Job, BAföG) zu erzielen – von dem sie sich aber nicht als grundsätzlich abhängig begreifen. Auch weisen Sie im Interview in der Regel darauf hin, dass für sie nicht die Maximierung des indivi-

duellen Auszahlungsbetrags, sondern die generelle positive Erfahrung der Experimentteilnahme im Vordergrund steht:

„[...] ICH FAND ES SEHR GUT, DASS WIR UNS SO SCHNELL IN DER ERSTEN RUNDE AUF DIESE VARIANTE D [GLEICHVERTEILUNG, DIE VERFASSER] GEEINIGT HABEN. DAS STELLT EINFACH EIN GERECHTES ERGEBNIS FÜR ALLE DAR UND ALLE KÖNNEN ZUFRIEDEN NACH HAUSE GEHEN.“

3.2.3 Typ 3: ‚Kritischer Experte‘

Im Gegensatz zum Narrativ der ‚Interessierten Forscherin‘, die sich aus einem gewissen professionellen (Fach-)Interesse beteiligt, nutzen die ‚Kritischen Experten‘ das Experiment, um – neben der finanziellen Entlohnung – einen Rahmen zu finden, ihr eigenes Handeln kritisch zu hinterfragen und zu reflektieren. Ähnlich der interessierten Forscher*innen betonen sie die thematische Ausrichtung des Experiments und äußern ihr genuin inhaltliches Interesse an wirtschaftlichen Verhalten:

„[...] WEIL ES MICH INTERESSIERT HAT, SO VERHALTEN IN DER WIRTSCHAFT. WIE LEUTE SICH VERHALTEN, WENN ES WIRKLICH UMS GELD GEHT. DAS IST JA NOCH WAS ANDERES, DA KOMMEN DANN JA WIRKLICH SO DIE – JA VIELLEICHT NICHT EGOISTISCHEN ABER DOCH SEHR EIGENNÜTZIGEN MOTIVE IRGENDWIE AUS DEN MENSCHEN RAUS. DARUM GING ES JA WAHRSCHEINLICH AUCH VERMUTE ICH MAL UNGEFÄHR“

Anders als die ‚Interessierten Forscher‘ nehmen die ‚Kritischen Expert*innen‘ eine kritische Haltung gegenüber der Gesellschaft ein, wobei diese Position von Selbstkritik und Hinterfragung des eigenen Verhaltens geprägt ist. Sie äußert, klar normativ orientiert, eine eindeutige Präferenz für eine „GERECHTE VERTEILUNG“. Kritische Expert*innen kritisieren entsprechend das „FEHLENDE POLITISCHE ENGAGEMENT“ ihrer Kommilitonen im Allgemeinen und speziell im durchgeführten experimentellen Setting. Im Interview betonen Sie ihre wechselhaften Gefühle und reflektieren ihren Umgang mit der experimentellen Situation:

„ES WAR HALT EIN RICHTIGES EXPERIMENT, KEIN VERSUCH ODER SOWAS UND ICH HATTE GANZ VERSCHIEDENE GEFÜHLE DABEI. AM ANFANG, ALS ICH DIE AUFGABENSTELLUNG ERFAHREN HABE UND SO WEITER HAB ICH GEDACHT: OH WIE BÖSE, SO VIEL KONKURRENZ UND ICH MAG DAS EIGENTLICH GAR NICHT. UND ((LACHT)) JA, ICH HAB AUCH ÄNGST GEHABT, DASS ICH SCHLECHT ABSCHNEIDE, WEIL FALLS ALSO AM ANFANG SCHON SO EINE VARIANTE MIT VIELEN STRATIFIKATIONEN GEWÄHLT WIRD, DASS ICH DANN IM ENDEFFEKT RAUSGEHE [...] ABER DANN HAB ICH MICH DRAUF EINGELASSEN UND GERADE NACH DEM ERSTEN KREUZWORTRÄTSEL, DAS ICH HALT FAST KOMPLETT GELÖST HATTE, HABE ICH DANN MICH AUCH RICHTIG GEFREUT UND DANN HAT ES MIR AUCH RICHTIG SPAß GEMACHT (LACHT). ALSO ES HAT MIR SPAß GEMACHT IM GROSSEN UND GANZEN. ICH FAND ES SPANNEND UND ICH HAB AUCH VIEL ÜBER MICH GELERNT DABEI“

Die Auseinandersetzung mit dem experimentellen Setting erzeugt eine Krise, welche mittels einer biographischen Verortung gelöst wird (Oevermann 1996): Die Proband*innen reflektieren ihren Umgang mit Konkurrenzerfahrungen und ihre eigene Einstellung im Experiment, legen aber zugleich Wert darauf, dass sie gleichermaßen wegen des Geldes gekommen sind. Da ungleiche Verteilungsmuster ihrer Identitätskonstruktion widersprechen und ihre Werthaltung berücksichtigen, präferieren sie – wie die ‚interessierten Forscher*innen‘ – egalitäre und nivellierende Verteilungsmuster. Sie sind „ÜBERRASCHT“ wie sehr das Experiment sie emotional „MITGENOMMEN“ hat und wie „LEICHT“ sie sich „ÜBER SO EINEN SCHEIß“ freuen können. Diese Erfahrung „IRRITIERT“ und ist den „Kritischen Expert*innen“ teilweise gar „EIN BISSCHEN PEINLICH“.

3.3 Verzerrung von Forschungsergebnissen durch nicht-repräsentatives Verhaltensweisen studentischer Experimentteilnehmer*innen

Es konnten drei Narrativtypen von studentischen Versuchsteilnehmer*innen identifiziert werden, die unseres Erachtens die Folge eines studentischen Samplings an Universitäten sind und der Anforderung nach externer Validität diametral entgegenstehen.⁶ Im Folgenden knüpfen wir an die allgemeinen Probleme der Rekrutierung studentischer Teilnehmer*innen aus Abschnitt 2.3 an und zeigen im Detail, dass unsere Befunde im Einklang mit der entsprechenden Literatur stehen. Die Teilnahmemotiken der drei identifizierten Typen lassen darauf schließen, dass eine mehrfache Teilnahme an Experimenten in ihrem jeweiligen spezifischen Interesse liegt: Der ‚Homo Oeconomicus‘-Typ verdient einen Teil seines studentischen Einkommens durch Studienteilnahmen, was oft als angenehmer und lukrativer empfunden wird als andere bezahlte Tätigkeiten. Die beiden anderen Typen lassen sich als Selbst- und Forschungserkenntnis-getriebene ‚passionierte Hobby-Teilnehmer*innen‘ beschreiben, die durch Mehrfachteilnahmen vor allem ihre Erkenntnis mehren wollen – die monetäre Entlohnung wird als willkommener Nebeneffekt eingestuft.

Die Existenz nicht-monetärer Teilnahmemotive ist indes in der Literatur schon länger bekannt. Neben den genannten Erkenntnisinteressen entschließen sich viele Proband*innen auch aus prosozialen Motiven heraus zu einer Anmeldung beim Proband*innenpool (Frey 1997; Charness/Rabin 2002; Fehr/Fischbacher 2002). Auch unsere anschließende Befragung ergab, dass sich die Teilnehmenden häufig untereinander kannten und wechselseitig auf das Experiment hingewiesen haben, also klare soziale Interaktionsmuster vorliegen. Man könnte fast von der Existenz einer *Experimental Subject Community* sprechen. Für die Durchführung von Experimenten ist es indes höchst problematisch, wenn es zu unkontrolliertem Austausch zwischen den Teilnehmenden vor dem Experiment kommt und ggf. bereits klar ist,

⁶ Eine Problematisierung, dass ausschließlich Studierende als Versuchspersonen auftreten, findet sich bereits bei Henrich et al. (2010).

wie das Experiment ungefähr abläuft. Ein solches *Priming* gilt es in Experimenten allgemein zu vermeiden, da es die interne Validität von Studien unterläuft; kann jedoch bei Existenz einer solchen ‚Community‘ keineswegs ausgeschlossen werden. Dass sich Teilnehmer*innen und Interessierte darüber austauschen, wo sich wie viel verdienen lässt, belegen auch Seiten wie *thepennyhoarder.com* (2014) oder *www.sowi-forum.com* (2008).⁷ Gerade letztere Seite verfügt über detaillierte Foreneinträge, die sich mit Entlohnung, Inhalten, und Verhaltenstipps in Experimenten beschäftigen. Letztlich werden so für den Experimentator unkontrolliert Informationen unter „Oft-an-Experimenten-Teilnehmer[n]“ ausgetauscht, die sowohl relevant für den monetär interessierten ‚Homo Oeconomicus‘-Typen sind, als auch die ‚Interessierte Forscherin‘ ansprechen. So schreibt beispielsweise User „volkswirt“ auf *www.sowi-forum.com* (2008):

„[o]ft geht es aber auch um public-goods-games, credence goods (Vertrauensgüter) und andere volkswirtschaftliche Probleme. Ist wirklich interessant!“

Dass Experimente insbesondere den ‚Homo Oeconomicus‘-Typen anziehen, zeigen die genannten Foreneinträge sowie unsere eigenen Daten – was natürlich nicht überrascht. In die gleiche Richtung deuten auch die Befunde von Krawczyk (2011) und Abeler/Nosenzo (2015): Mittels Modifikation der Anschreiben zeigten diese Studien, dass die Teilnahmewahrscheinlichkeit ohne monetäre Entlohnung deutlich fällt. Guillén/Veszteg (2012) gemäß erhöht eine vorherige positive Auszahlung in Experimenten die zukünftige Teilnahmewahrscheinlichkeit signifikant. Das spricht für einen erheblichen Anteil an ‚Homo Oeconomicus‘-Teilnehmer*innen, der auch in unserem Sample immerhin fast 60% der Teilnehmerschaft umfasste.

Aber auch der von uns identifizierte Typ ‚Interessierter Forscher‘ passt zu den Beobachtungen anderer Autoren, die die begründete Vermutung äußern, dass sich „eher interessierte ‚scientific do-gooders‘“ für Experimente anmelden (Berger 2015: 57; Ball/Cech 1996: 241; Levitt/List 2007: 165; ähnlich bereits Rosenthal/Rosnow 1975).

Ein Problem, das Typen-übergreifend auftritt, ist das der Vielfachteilnahme. Dass sich Vielfach-Teilnehmer*innen an einem Experimental-Standort auch ohne ein Online-Forum intensiv austauschen und ohne dass es von der Studienleitung unterbunden werden kann, erscheint sehr plausibel. Insgesamt sorgt dies für eine ‚Artifizialisierung‘ des Verhaltens von Teilnehmern, wobei sich die konkret entstehenden Verzerrungen unserer Ansicht nach typenspezifisch darstellen. Berger hat darauf hingewiesen, dass insbesondere an Standorten, die bereits länger und in größerem Umfang Laborexperimente durchführen, die Proband*innenpopulation dadurch verzerrt wird, dass bestimmte Proband*innen(gruppen) öfter an Experimenten teil-

7 Auch wenn diese Seite seit zehn Jahren nicht mehr genutzt wird, spricht nichts gegen unsere Vermutung. Vielmehr ist davon auszugehen, dass andere Kommunikationsformen wie WhatsApp- oder Facebook-Gruppen, die von Experimentator*innen schwieriger einzusehen sind, solche offenen Foren verdrängt haben.

nehmen als andere (Berger 2015: 57). Solche „professionellen Probanden“ (Berger 2015: 57), die weitgehend dem Verhaltenstyp ‚Homo Oeconomicus‘ entsprechen, unterscheiden sich deutlich von der ursprünglichen Population und zeigen aufgrund ihrer Erfahrungen und Lernkurven im Umgang mit Laborexperimenten deutlich andere Verhaltensweisen (Bardsley 2005; Levitt/List 2007; Benz/Meier 2008).

Aufgrund unserer Befunde ist insbesondere bei ökonomischen Verteilungsexperimenten auf die Gewinnmaximierungsstrategie durch mehrmalige Teilnahme bei gleichzeitiger höherer Risikoneigung hinzuweisen. Durch die Vielfachteilnahme taucht hier ein Problem auf, dass sich auch durch Kontrolle eines eigenen Proband*innenpools nicht einfach in den Griff bekommen wird: Studierende nehmen nicht nur an Experimenten teil, die durch Ökonom*innen angeboten und organisiert werden, sondern je nach Standort auch an medizinischen, psychologischen usw. Studien. Dabei streuen gewinnmaximierende ‚Homo Oeconomicus‘-Typen – ganz rational – ihr Entscheidungsrisiko unkontrolliert über verschiedene Experimente. Wer beispielsweise bereits am Nachmittag an der nächsten Studie teilnimmt, kann im Experiment am Vormittag durchaus eine Strategie wählen, die zwar ggf. zu einem niedrigeren Auszahlungswert führt, aber eben auch die Chance auf einen hohen Gewinn mit sich bringt. Der fix bezahlte medizinische Versuch am Nachmittag fungiert – ohne, dass es der Experimentatorin am Vormittag bewusst ist – als Absicherung (bzw. als ‚Mindesteinkommen‘) für den Fall, dass der Gewinnmaximierer im einzelnen Experiment ‚Pech‘ hat.

Gerade in ökonomischen Experimenten, in denen Risikobereitschaft von zentraler Bedeutung ist, unterlaufen ‚Experiment-Hedger‘ die Verlässlichkeit der Messungen, ganz abgesehen von der unbeobachteten grundsätzlichen ‚Professionalisierung‘ der Proband*innen durch Teilnahme an anderen Studien. Im Prinzip müssten also Standorte, in denen Studien in unterschiedlichen Fachrichtungen durchgeführt werden, einen gemeinsamen Pool einrichten und Experimententeilnehmer*innen wechselseitig ausschließen. Aus logistischen Gründen ist dies aber oftmals nicht der Fall, zudem ist damit nicht gesichert, dass ‚professionelle‘ Studienteilnehmer*innen zudem an außeruniversitären Versuchen, anderen Universitäten etc. teilnehmen, was letztlich aber einen ähnlich verzerrenden Effekt auf das Verhalten in einem konkreten ökonomischen Experiment hat.

Auch für die ‚Interessierte Forscherin‘ ist es sinnvoll, an möglichst vielen Studien teilzunehmen, um ihr Erkenntnisinteresse zu befriedigen. Je enger dabei das Interessensgebiet, desto wahrscheinlicher ist es, dass eine Person nur ganz spezifische Experimente aufsucht und eine entsprechend große Expertise mitbringt, deren Umfang sich bis zu einem erheblichen *Priming* auswachsen kann. Und selbst bei einem recht allgemeinen Interesse an ‚experimentellen‘ oder ‚empirischen‘ Methoden verändert sich das Verhalten aufgrund bestehender Vorerfahrung wie Bardsley (2005), Levitt/List (2007) und Benz/Meier (2008) zeigen. Gewöhnlich reagieren Labore auf solche

Probleme, indem Personen nicht mehr zu bestimmten Experimenten eingeladen werden, wenn sie vorher an einem ähnlichen Experiment teilgenommen haben. Zudem werden ‚verbrauchte‘ Proband*innen nach einer gewissen Zahl an Studienteilnahmen aus dem Verteiler gelöscht. Allerdings kann die Teilnahme an Studien anderer Einrichtungen nur schwierig kontrolliert werden, wenn Psychologie- und Ökonomik-Labore an einem Universität-Standort koexistieren. Natürlich ist es für ‚Interessierte Forscher*innen‘ und ‚Kritische Expert*innen‘ schwieriger als für den ‚Homo Oeconomicus‘, an möglichst vielen Experimenten teilzunehmen: Finanzielle Anreize werden (fast) überall geboten, aber nicht jedes monetär entlohnte Experiment mag ins eigene Fachgebiet fallen oder eine ethische Thematik ansprechen. Zudem wird die Vielfachteilnahme durch das übliche Herausfiltern von Vielfach-Teilnehmer*innen in spezifischen Umfeldern erschwert.

Dennoch ergeben sich auch für diese beiden Typen artifizielle Experimentvorteile. Die ‚Interessierte Forscherin‘ hat den Vorteil, dass sie sich im eigenen Studium intensiv in ein Thema einlesen kann und all dieses Wissen – etwa um spieltheoretisch optimale Entscheidungen – in ihr Verhalten einfließen lassen kann. Für den ‚Kritischen Experten‘ gilt, ähnlich wie für die ‚Interessierte Forscherin‘, dass Vielfachteilnahmen die Erkenntnisbasis erweitern und daher attraktiv erscheinen. Zudem besteht auch hier der Vorteil des Vorab-Zugangs mittels Literatur oder Selbsterfahrung, wenn das Erkenntnisinteresse – wie oben beschrieben – anders gelagert ist. Für alle drei Typen gilt somit gleichermaßen, dass sie sich als Vielfachteilnehmer sehr professionell und damit artifiziell verhalten.

Insgesamt lässt sich festhalten: Die Selbstelektion bestimmter Typen ins Experiment unterläuft die externe Validität, da Effekte nur für diese spezielle Population der *experimental community* gemessen werden. Mit Slonim et al. (2013) folgern wir, dass zu bezweifeln ist, ob damit belastbare Aussagen über das Verhalten der Zielpopulation – Konsument*innen, Bürger*innen, Investor*innen etc. – möglich sind. Das sehr trainierte und einer eigenen Interessenlogik folgende Verhalten der studentischen Teilnehmer*innen legt zudem nahe, dass die Proband*innen das Experiment vielfach völlig anders interpretieren, als von den Experimentatoren beabsichtigt. Dies unterläuft letztlich die interne Validität, da nicht klar ist, was nun genau gemessen wird – eine direkte Folge des nicht-repräsentativen, artifiziellen Proband*innen-Pools.

4 Diskussion und Empfehlungen für soziologische Experimente

Zusammenfassend ist festhalten, dass die zentrale Kritik an der experimentellen Sozialforschung sich auf die mangelnde Übertragbarkeit der experimentell gewonnenen Erkenntnisse auf die Welt außerhalb des Labors bezieht (externe Validität). Weitere Kritikpunkte umfassen u.a. die permanente Wiederholung bestimmter Verhaltenssituationen, die fehlende Motivation der Akteur*innen sowie die Zusammensetzung der Versuchsgruppen, die meist ausschließlich durch Student*innen

gebildet werden. Diese weisen andere Verhaltensweisen auf als Entscheidungsträger*innen in realen sozialen Interaktionen. Schon die Bereitschaft zur Teilnahme an einem Experiment kann Indikator für überdurchschnittlich kooperatives Verhalten sein. Die genannten Probleme rühren Großteils daher, dass im Experiment ein möglichst hoher Grad an interner Validität erreicht werden soll: Studentische Versuchsgruppen sind, im Vergleich zur Gesamtbevölkerung, relativ homogen, so dass sich Treatment-Effekte klarer identifizieren lassen. Ein Effekt – genauso wie sein Ausbleiben – mag allerdings am Ende der artifiziellen, untypisch agierenden Population geschuldet sein. Unser Beitrag geht über diese bekannte Kritik an der Verwendung studentischer Proband*innen-Pools hinaus: Die Selbstselektion der Proband*innen *in den Pool* sorgt dafür, dass die Proband*innen nicht einmal repräsentativ für die studentische Gesamtpopulation sind. Entgegen der unter experimentellen Forscher*innen verbreiteten Ansicht, (Nicht-)Repräsentativität sei kein Problem für die Untersuchung von Kausalzusammenhängen, sehen wir dieses Problem äußerst kritisch.

Ein Experiment erlaubt stets nur Aussagen über den Effekt in der untersuchten Gruppe. Größe oder gar Richtung des Effekts mögen in einer anderen Gruppe mit anderem sozioökonomischen Hintergrund signifikant anders aussehen (Slonim et al. 2013). Zwar könnte argumentiert werden, dass diese Verzerrung für die interne Validität des Experiments keine Relevanz besitzt, weil Experimentator*innen eine möglichst repräsentative Gruppe aus dem vorhandenen Proband*innen-Pool auswählen und die Teilnehmerschaft dann randomisiert in Kontroll- und Treatment-Gruppe aufteilen. Dennoch bleibt das Problem bestehen, dass sich der Proband*innenpool durch Selbstselektion gebildet hat und relevante, unbeobachtete Charakteristika (wie Teilnahmemotivation, Erfahrungen mit Experimenten, soziale Beziehungen zu anderen Experimententeilnehmer*innen etc.) nicht kontrolliert werden. Zusammengefasst: Selbstselektion findet nicht *im* Experiment statt, dafür aber *ins* Experiment – mit der genannten Folge, dass die beobachteten Effekte erst einmal nur für diese selbstselektierte Population gelten. Die oftmals *a priori* postulierte Verallgemeinerbarkeit müsste dann erst gezeigt werden.

Insgesamt können Laborexperimente also nur bedingt reale Verhaltenssituationen approximieren, weshalb ihre Ergebnisse mit Vorsicht zu behandeln sind (vgl. Schoefer 2005: 33; Konow 2003: 1191-1193). Die in diesem Beitrag identifizierte Typologie verweist auf ein zentrales Problem sozialwissenschaftlicher Experimente: Handlungsmotivation und Handlungsmuster unterscheiden sich je nach Handlungstyp. Zudem kennen sich die Versuchsteilnehmer*innen häufig, sind mit Experimenten vertraut und modifizieren entsprechend ihr Verhalten, weshalb zu erwarten ist, dass Kooperationsraten „positiv verzerrt, nicht reliabel und deshalb auch nicht valide sind“ (Berger 2015: 68).

Üblicherweise wird empfohlen, Experimente mit anderen Teilnehmer*innen und an anderen Standorten zu replizieren, die Experimente mit denselben Teilneh-

mer*innen unter Kontrolle möglicher Lerneffekte zu einem späteren Zeitpunkt zu wiederholen und auf instituts- oder universitätsübergreifende Proband*innenpools⁸ zu verzichten (Berger 2015: 73). Während wir den ersten beiden Punkten zustimmen, sehen wir bezüglich der Teilnehmendenpools erheblichen Diskussionsbedarf. Der Verzicht auf übergreifende Proband*innenpools bedeutet nicht unbedingt eine größere Kontrolle über die Teilnehmer*innen, da der Austausch über vorherige Studienteilnahmen an unterschiedlichen Standorten verloren geht. Vor diesem Hintergrund gilt es kritisch hervorzuheben, dass die weithin bekannten Probleme und Handlungsalternativen „zu wenig beachtet“ werden (Berger 2015: 71). In der Diskussion überrascht daher, dass ein zentrales Problem experimenteller Sozialforschung – die überwiegende Rekrutierung aus der Studierendenschaft – in der Praxis nicht stärker diskutiert wird (vgl. exemplarisch Abeler/Nosenzo 2015 oder die Beiträge in Keuschnigg/Wolbring 2015). Die Problematik des *Convenience Samplings* aufgrund einfacher Verfügbarkeit von Studierenden wird zwar benannt, eine Abkehr von dieser Praxis wird jedoch weder systematisch gefordert noch umfassend praktiziert.

Wie gezeigt sind Studierende aus zwei Gründen problematische Proband*innen. Erstens stammen Studierende überdurchschnittlich oft bildungsnahen und sozial höher gestellten Schichten (BMBF 2004) und zweitens sind sie selbst Teil des Analysefeldes und werden im Laufe ihrer Ausbildung zu ‚Expert*innen‘. Entsprechend wird gefordert, dass zukünftig den methodischen Problemen der Stichprobenziehung sowie der Selektivität und Abnutzung von Stichproben mehr Aufmerksamkeit geschenkt werden müsse (Becker/Zangger 2015: 113). Die Selektivität der Proband*innengruppen hat möglicherweise zu allerlei Artefakten bei den statistischen Analysen geführt (Henrich et al. 2010; Slonim et al. 2013).

Die Vermeidung der genannten Probleme ließe sich am besten durch Verzicht auf studentische Samples und die Nutzung repräsentativer Stichproben erreichen. Dies scheint uns allerdings unrealistisch, da Forschung unter finanziellem und zeitlichem Druck stattfindet. Daher schlagen wir folgende Verbesserungsmöglichkeiten vor, die bei studentischen Stichproben die Aussagekraft des Experiments erhöhen würden.

Ein erster Schritt wäre – im expliziter Abgrenzung zu dem Vorschlag von Berger – die Erstellung von kontrollierten Proband*innenpools, optimalerweise über verschiedene Fachdisziplinen und Standorte hinweg, um zumindest die Gefahr einer nicht erwünschten Mehrfachteilnahme zu verhindern und die verzerrenden Lerneffekte zu unterdrücken. Hier würde sich ein anonymisiertes Identifizierungssystem für Experimente anbieten, welches jedem Studierenden in Deutschland einen Code

8 Das bekannte Vorgehen, dass ein Labor einen eigenen Pool unterhält und zumindest innerhalb diesem Teilnehmer*innen kontrolliert, ist uns geläufig und ist natürlich methodisch absolut geboten. Wie oben beschrieben plädieren wir aber für eine Fortentwicklung dieses Vorgehens, um Teilnahme-‚Externalitäten‘ und damit Verzerrungen von Laborergebnissen weiter zu mindern.

zur Teilnahme an Experimenten nach Anmeldung bzw. Registrierung mittels Matrikelnummer zur Verfügung stellt.⁹ Darüber hinaus sollten vorherige Teilnahmen an Experimenten abgefragt und analytisch berücksichtigt werden. Zu diskutieren wäre, ob potenzielle ‚Vielexperimentierer‘ für weitere Teilnahmen gesperrt werden sollten.

Ein zweiter Schritt wäre, studentische Samples stärker zu kontrollieren. Wir plädieren für eine umfassende Studie zu den Selektionsmerkmalen von Experimentteilnehmer*innen mit Blick auf Fachrichtungen, ethnischer Herkunft, Geschlecht, sozioökonomische Hintergründe etc. Unsere Befunde zeigen, dass nicht nur die Auswahl *ex ante* kontrolliert werden muss sondern gleichermaßen eine *ex post*-Prüfung nötig ist. Nur so kann man einen Überblick über die tatsächlichen Verzerrungen studentischer Samples erlangen. Zudem wäre zu fragen, ob sich die identifizierten Teilnahmetypen an verschiedenen anderen Labor-Standorten und Experimenten ebenfalls so finden lassen und über welche Häufigkeiten sich die Typen verteilen. Wenn bekannt ist, wie sich die identifizierten Typen in der Gesamtbevölkerung (oder zumindest der studentischen Population) verteilen, ließen sich die im Experiment vorgefundenen Typenverteilungen repräsentativ umgewichten.¹⁰ Die Identifizierung solcher Typen würde zunächst aufwändige qualitative Verfahren wie unsere geführten Interviews erfordern. Sobald aber ausreichend Material zur Typenbildung vorhanden ist, sollte die Zuordnung von Studienteilnehmer*innen durch standardisierte Verfahren (z.B. Fragebatterien) erfolgen. Der typenbasierte Ansatz zur Erreichung (oder zumindest Approximierung) von Vergleichbarkeit scheint uns angesichts der unterschiedlichen typenspezifischen Verhaltensmuster eine sinnvolle, gar notwendige qualitative Erweiterung der quantitativen Logik, Repräsentativität allein durch statistische (Mittelwert-)Verfahren zu erreichen.

Ein weiteres Problem sozialwissenschaftlicher Ansätze, die vom beobachtbaren Verhalten auf die zugrundeliegende Motivation schließen, ist, dass die Entscheidenden selbst keine Möglichkeit haben, ihre Beweggründe für ihre Entscheidung zu erklären. Das Argument von beispielsweise (Experimental-)Ökonom*innen (Wardman 1988, Brown 2003), dass nur quantifizierbare Entscheidungen verlässlich Aufschluss über individuelle Beweggründe geben, überzeugt nicht: Erstens sind experimentelle Entscheidungssituationen selten derart gestaltet, dass eine Entscheidung nur durch ein einziges, isolierbares Motiv erklärbar ist. Zweitens ist der Vorwurf, Proband*innen berichteten in Interviews häufig nicht die Wahrheit, gerade mit Blick auf Erkenntnisse aus der Verhaltensökonomik nicht pauschal haltbar: Aktiv die Unwahrheit zu sagen ist auf Dauer recht anstrengend (Brown 2003, Sánchez-Pagés/Vorsatz 2009). Gewonnen werden durch Interviews (oder wenigstens offene schriftliche Verhaltensbegründungen) aber zumindest Einblicke in die Motivation

9 Wobei entsprechende datenschutzrechtliche Probleme noch zu diskutieren wären.

10 Offen ist, wie mit dem Problem umzugehen wäre, dass bestimmte Typen nur in der studentischen Population vorkommen, nicht aber (oder nur kaum) in der Gesamtbevölkerung.

der Teilnehmer*innen, die bei der weiteren Theorie- und Hypothesenbildung Berücksichtigung finden müssen.

Dies führt wieder zum Ausgangsthema zurück: Neben der Typisierung hilft ein solches Instrument, zwischen Entscheidungsgründen zu unterscheiden, die dem Experiment zuzuordnen sind, und solchen, die mit hoher Wahrscheinlichkeit auf individuelle, habituelle Verhaltensmuster schließen lassen, welche von der Lebenswelt ins Labor übertragen wurden. So belegen unsere Interviews den skizzierten Motivmix, wobei sich der Grad an Anpassung an ein Laborsetting typenabhängig zeigte. Damit lässt sich auch der Grad der externen Validität besser beurteilen. Entscheidend ist letztlich nicht, ob die durchführenden Wissenschaftler*innen den Versuch als ‚ausreichend realitätsnah‘ betrachten bzw. die wissenschaftlichen Peers zu diesem Schluss kommen, sondern wie die Proband*innen das Setting subjektiv interpretieren. Um dies zu beurteilen, können und müssen die Teilnehmer*innen unseres Erachtens nach zwingend hinsichtlich ihrer Motivlage befragt bzw. ihre Interpretation der Versuchsanordnung erhoben werden.

Insgesamt sind wir überzeugt, dass dem Problem der ‚professionellen‘ bzw. ‚Hobby-Proband*innen‘ begegnet werden muss. Einige mögliche Abhilfen haben wir in diesem Beitrag skizziert. Wir hoffen, dass unsere oben skizzierten Anregungen in der experimentellen Sozialforschung aufgenommen werden, da die experimentelle Methode in den Sozialwissenschaften sehr wohl ihre Berechtigung hat und Ergebnisse produziert, die durch andere Methoden so nicht erreicht werden können. Um diesen Mehrwert aufrechtzuerhalten, müssen die beschriebenen Probleme allerdings ernsthaft angegangen werden, damit die Vorteile der experimentellen Methode nicht verspielt werden und der Ansatz hinter die Leistungen von Surveymethoden zurückfällt.

Literaturverzeichnis

- Abeler, Johannes / Nosenzo, Daniele (2015): Self-Selection into Laboratory Experiments: Pro-Social Motives Versus Monetary Incentives, in: *Experimental Economics* 18, S. 195-214.
- Albert, Hans (1965): Modell-Platonismus. Der neoklassische Stil des ökonomischen Denkens in kritischer Beleuchtung, in: Ernst Topitsch (Hg.): *Logik der Sozialwissenschaften*. Berlin, S. 406-434.
- Arnett, Jeffrey J. (2008): The Neglected 95%. Why American Psychology Needs to Become Less American, in: *American Psychologist* 63, S. 602-614.
- Aronson, Elliot / Wilson, Timothy D. / Akert, Robin M. (2004): Sozialpsychologie. 4., aktualisierte Aufl. München.
- Bahr, Gunter / Requate, Till (2013): Reciprocity and Giving in a Consecutive Three-Person Dictator Game with Social Interaction, in: *German Economic Review* 15, 374-392.
- Bailey, Phoebe E. / Ruffman, Ted / Rendell, Peter G. (2013): Age-Related Differences in Social Economic Decision Making: The Ultimatum Game, in: *Journal of Gerontology B Series* 68, S. 356-363.

- Ball, Sheryl B. / Cech, Paula-Ann (1996): Subject Pool Choice and Treatment Effects in Economic Laboratory Research, *Research in Experimental Economics* 6, S. 239-292.
- Bardsley, Nicholas (2005): Experimental Economics and the Artificiality of Alteration, in: *Journal of Economic Methodology* 12, S. 239-251.
- Becker, Rolf / Zangger, Christoph (2015): Experimentelle Bildungsforschung. Beiträge experimenteller Forschung zur Bildungssoziologie und -ökonomie, in: Marc Keuschnigg / Tobias Wolbring (Hg.): *Experimente in den Sozialwissenschaften*. Soziale Welt, Sonderband 22. Baden-Baden, S. 105-131.
- Benz, Matthias / Meier, Stephan (2008): Do People Behave in Experiments as in the Field? Evidence from Donations, in: *Experimental Economics* 11, S. 268-281.
- Berger, Joël / Diekmann, Andreas (2015): The Logic of Relative Frustration: Boudon's Competition Model and Experimental Evidence, in: *European Sociological Review* 31, S. 725-737.
- Berger, Roger (2015): Das Laborexperiment als sozialer Prozess, in: Marc Keuschnigg / Tobias Wolbring (Hg.): *Experimente in den Sozialwissenschaften*. Soziale Welt, Sonderband 22. Baden-Baden, S. 53-76.
- Binmore, Ken / Shaked, Avner (2010): Experimental Economics: Where Next?, in: *Journal of Economic Behavior and Organization* 73, S. 87-100.
- Brown, Thomas C. (2003): Introduction to Stated Preference Methods, in: Patricia A. Champ / Kevin J. Boyle / Thomas C. Brown (Hg.): *A Primer on Nonmarket Valuation*. Dordrecht, Boston, S. 99-110.
- Bundesministerium für Bildung und Forschung (BMBF) (2004): Die wirtschaftliche und die soziale Lage der Studierenden in der Bundesrepublik Deutschland 2003, Berlin.
- Burger, Jerry M. (2009): Replicating Milgram. Would People Still Obey Today?, in: *American Psychologist* 64, S. 1-11.
- Camerer, Colin F. (2003): Behavioral Game Theory. Experiments in Strategic Interaction, Princeton.
- Carpenter, Jeffrey P. / Harrison, Glenn W. / List, John A. (2005): Field Experiments in Economics: An Introduction, in: *Research in Experimental Economics* 10, S. 1-16.
- Charness, Gary / Rabin, Matthew (2002): Understanding Social Preferences with Simple Tests, in: *Quarterly Journal of Economics* 117, S. 817-869.
- Chuah, Swee-Hoon / Hoffmann, Robert / Jones, Martin / Williams, Geoffrey (2007): Do Cultures Clash? Evidence From Cross-National Ultimatum Game Experiments, in: *Journal of Economic Behavior & Organization* 64, S. 35-48.
- Cleave, B. L. / Nikiforakis, N. / Slonim, R. (2013): Is There Selection Bias in Laboratory Experiments? The Case of Social and Risk Preferences, in: *Experimental Economics* 16, S. 372-382.
- Cook, Thomas D. / Campbell, Donald T. (1979): Quasi-Experimentation. Design and Analysis Issues for Field Settings, Boston.
- Croson, Rachel (1996): Information in Ultimatum Games: An Experimental Study, in: *Journal of Economic Behavior and Organization* 30, S. 197-212.
- Croson, Rachel / Gächter, Simon (2010): The Science of Experimental Economics, in: *Journal of Economic Behavior & Organization* 73, S. 122-131.
- Davis, Douglas D. / Holt, Charles A. (1993): Experimental Economics, Princeton.

- Diekmann, Andreas (2008): Soziologie und Ökonomie: Der Beitrag experimenteller Wirtschaftsforschung zur Sozialtheorie, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 60, S. 528-550.
- Diekmann, Andreas (2015): Geleitwort, in: Marc Keuschnigg / Tobias Wolbring (Hg.): *Experimente in den Sozialwissenschaften*. Soziale Welt, Sonderband 22. Baden-Baden, S. 3-5.
- Falk, Armin (2001): Wirtschaftswissenschaftliche Experimente: Homo Oeconomicus auf dem Prüfstand, in: *Wirtschaftsdienst – Zeitschrift für Wirtschaftspolitik* 81, S. 300-304.
- Falk, Armin / Meier, Stephen / Zehnder, C. (2013): Do Lab Experiments Misrepresent Social Preferences? The Case of Self-Selected Student Samples, in: *Journal of the European Economic Association* 11, S. 839-852.
- Fehr, Ernst / Fischbacher, Urs (2002): Why Social Preferences Matter. The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives, in: *The Economic Journal* 112, C1-C33.
- Fehr, Ernst / Schmidt, Klaus M. (1999): A Theory of Fairness, Competition and Cooperation, in: *Quarterly Journal of Economics* 114, S. 817-868.
- Frey, Bruno S. (1997): *Not Just for the Money. An Economic Theory of Personal Motivation*, Cheltenham.
- Friedman, Daniel / Cassar, Alessandra (2004): *Economics Lab. An Introduction to Experimental Economics*, London.
- Friedman, Daniel / Sunder, Shyam (1994): *Experimental Methods. A Primer for Economists*, Cambridge.
- García-Gallego, Aurora / Georgantzís, Nikolaos / Jaramillo-Gutiérrez, Ainhoa (2012): Gender Differences in Ultimatum Games. Despite Rather than Due to Risk Attitudes, in: *Journal of Economic Behavior & Organization* 83, S. 42-49.
- Guillén, Pablo / Veszteg, Róbert F. (2012): On "Lab Rats", in: *The Journal of Socio-Economics* 41, S. 714-720.
- Güth, Werner / Kocher, Martin G. (2014): More than Thirty Years of Ultimatum Bargaining Experiments: Motives, Variations, and a Survey of the Recent Literature, in: *Journal of Economic Behavior & Organization* 108, S. 396-409.
- Güth, Werner / Schmittberger, Rolf / Schwarze, Bernd (1982): An Experimental Analysis of Ultimatum Bargaining, in: *Journal of Economic Behavior and Organization* 3, S. 367-388.
- Güth, Werner / van Damme, Eric (1998): Information, Strategic Behavior, and Fairness in Ultimatum Bargaining: An Experimental Study, in: *Journal of Mathematical Psychology* 42, S. 227-247.
- Harbaugh, William T. / Krause, Kate / Vesterlunge, Lise (2007): Learning to Bargain, in: *Journal of Economic Psychology* 28, S. 127-142.
- Harrison, Glenn W. / List, John A. (2004): Field Experiments, in: *Journal of Economic Literature* 42, S. 1009-1055.
- Held, Martin / Kubon-Gilke, Gisela / Sturn, Richard (Hg.) (2003): *Experimente in der Ökonomik*, Marburg.
- Helfferich, Cornelia (2011): *Die Qualität qualitativer Daten. Manual für die Durchführung qualitativer Interviews*. 4. Aufl., Wiesbaden.
- Henrich, Joseph (2000): Does Culture Matter in Economic Behavior? Ultimatum Game Bargaining among the Machiguenga of the Peruvian Amazon, in: *American Economic Review* 90, S. 973-979.

- Henrich, Joseph / Boyd, Robert / Bowles, Samuel / Camerer, Colin / Fehr, Ernst / Gintis, Herbert / McElreath, Richard (2001): In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies", in: *American Economic Review* 91, S. 74-78.
- Henrich, Joseph / et al. (2005): 'Economic Man' in Cross-Cultural Perspective: Behavioral Experiments in 15 Small-Scale Societies, in: *Behavioral and Brain Sciences* 28, S. 1-61.
- Henrich, Joseph / Heine, Steven J. / Norenzayan, Ara (2010): The Weirdest People in the World?, in: *Behavioral and Brain Sciences* 33, S. 61-83.
- Hoffmann, Robert / Tee, Jin-Yee (2006): Adolescent-Adult Interactions and Culture in the Ultimatum Game, in: *Journal of Economic Psychology* 27, S. 98-116.
- Jackson, M. / Cox, D. R. (2013): The Principles of Experimental Design and Their Application, in: *Annual Review of Sociology* 39, S. 27-49.
- Kahneman, Daniel / Knetsch, Jack L. / Thaler, Richard (1986): Fairness as a Constraint on Profit Seeking: Entitlements in the Market, in: *American Economic Review* 76, S. 728-741.
- Kelle, Udo / Kluge, Susanne (1999): Vom Einzelfall zum Typus. Fallvergleich und Fallkontrastierung in der qualitativen Sozialforschung, Opladen.
- Keuschnigg, Marc / Wolbring, Tobias (Hg.) (2014): Themenheft Experimente, in: *Soziale Welt* 2/14, Baden-Baden.
- Keuschnigg, Marc / Wolbring, Tobias (2015 a): Einleitung, in: Marc Keuschnigg / Tobias Wolbring (Hg.): *Experimente in den Sozialwissenschaften*. Soziale Welt, Sonderband 22. Baden-Baden, S. 7-13.
- Keuschnigg, Marc / Wolbring, Tobias (Hg.) (2015 b): Experimente in den Sozialwissenschaften, Soziale Welt Sonderband 22, Baden-Baden.
- Kleining, Gerhard (1982): Umriss zu einer Methodologie qualitativer Sozialforschung, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 34, S. 224-253.
- Konow, James (2003): Which Is the Fairest One of All? A Positive Analysis of Justice Theories, in: *Journal of Economic Literature* 41, S. 1188-1239.
- Krawczyk, Michal (2011): What Brings Your Subjects to the Lab? A Field Experiment, in: *Experimental Economics* 14, S. 482-489.
- Kruse, Jan (2014): Qualitative Interviewforschung. Ein integrativer Ansatz, Weinheim.
- Kuckartz, Udo / Dresing, Thorsten / Rädiker, Stefan / Stefer, Claus (2008): Qualitative Evaluation – Der Einstieg in die Praxis. 2., aktualisierte Auflage, Wiesbaden.
- Levitt, Steven D. / List, John A. (2007): What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?, in: *Journal of Economic Perspectives* 21, S. 153-174.
- Levitt, Steven D. / List, John A. (2009): Field Experiments in Economics: The Past, the Present, and the Future, in: *European Economic Review* 53, S. 1-18.
- Lorenz, Stephan / Gross, Matthias (2010): Wissen, Experimente, Verfahren. Zur Methodologisierung von Unsicherheit, in: Sozialwissenschaftlicher Fachinformationsdienst (Hg.): *Methoden und Instrumente der Sozialwissenschaften*. Bonn, S. 11-18.
- Mayring, Philipp (2015): Qualitative Inhaltsanalyse. Grundlagen und Techniken. 12., überarbeitete Auflage, Weinheim.
- Milgram, Stanley (1974): Das Milgram-Experiment. Zur Gehorsamsbereitschaft gegenüber Autorität, Reinbek bei Hamburg.
- Nowak, Martin A. / Page, Karen M. / Sigmund, Karl (2000): Fairness versus Reason in the Ultimatum Game, in: *Science* 5485, S. 1773-1775.

- Oevermann, Ulrich (1996): Krise und Muße. Struktureigenschaften ästhetischer Erfahrung aus soziologischer Sicht. Vortrag am 19.6. in der Städel-Schule. Online verfügbar unter publikationen.ub.uni-frankfurt.de/files/4953/Krise-und-Musse-1996.pdf, zuletzt geprüft am 29. September 2015.
- Oosterbeek, Hessel / Sloof, Radolph, Van den Kuilen (2004): Cultural Differences in Ultimatum Game Experiments: Evidence from a Meta-Analysis, in: *Experimental Economics* 7, S. 171-188.
- Plott, Charles R. (1991): Will Economics Become an Experimental Science?, in: *Southern Economic Journal* 57, S. 901-919.
- Roe, Brian E. / Just, David R. (2009): Internal and External Validity in Economic Research: Tradeoffs between Experiments, Field Experiments, Natural Experiments, and Field Data, in: *American Agricultural Economics Journal* 91, S. 1266-1271.
- Rosenthal, Robert / Rosnow, Ralph L. (1975): *The Volunteer Subject*, New York.
- Saad, Gad / Gill, Tripat (2001): Sex Differences in the Ultimatum Game: An Evolutionary Psychological Perspective, in: *Journal of Bioeconomics* 3, S. 171-193.
- Salganik / Matthew J. / Dodds, Peter Sheridan / Watts, Duncan J. (2006): Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market, in: *Science* 311, S. 854-856.
- Sánchez-Pagés, Santiago / Vorsatz, Marc (2009): Enjoy the Silence: An Experiment on Truth-Telling, in: *Experimental Economics* 12, S. 200-241.
- Schnell, Rainer / Hill, Paul B. / Esser, Elke (1999): *Methoden der empirischen Sozialforschung*. 6. völlig überarb. und erw. Aufl., München.
- Schoefer, Martin (2005): *Ökonomik – experimentelle Wirtschaftsforschung – Wirtschaftsethik*, Münster.
- Schröder, Martin (2011): Vom Experiment zur Praxis: Wie moralische Argumente wirtschaftliche Selbstinteressen beeinflussen, *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 63, S. 61-81.
- Shadish, William R. / Cook, Thomas D. / Campbell, Donald T. (2002): *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Boston.
- Slonim, Robert / Wang, Carmen / Garbarino, Ellen / Merrett, Danielle (2013): Opting-In: Participation Bias in Economic Experiments, in: *Journal of Economic Behavior & Organization* 90, S. 43-70.
- Smith, Vernon L. (2008): Experimental Methods in Economics, in: Steven N. Durlauf / Lawrence E. Blume (Hg.): *The New Palgrave Dictionary of Economics*. 2nd Edition. Houndmills.
- Solnick, Sara J. (2001): Gender Differences in the Ultimatum Game, in: *Economic Inquiry* 39, S. 189-200.
- Sowi-Forum (2008). Thema: Ökonomische Experimente: Ein einfacher Nebenverdienst, URL: <https://www.sowi-forum.com/forum/threads/30767-%C3%96konomische-Experimente-Ein-einfacher-Nebenverdienst/page2> [19.01.2018]
- The Penny Hoarder (2014). Get Paid to Be a Test Subject: 5 Ways to Earn Money as a Human Guinea Pig, URL: <https://www.thepennyhoarder.com/make-money/side-gigs/get-paid-test-subject-5-ways-earn-money-human-guinea-pig> [19.01.2018]
- Tisserand, Jean-Christian (2016): Do they Act Rationally? The Ultimatum Game, a Meta-Analysis of Three Decades of Experimental Research, in: *Empirical Economics Letters* 15.
- Wardman, Mark (1988): A Comparison of Revealed Preference and Stated Preference Models of Travel Behaviour, in: *Journal of Transport Economics and Policy* 22, S. 71-91.

Webster, Murray / Sell, Jane (Hg.) (2014): Laboratory Experiments in the Social Sciences. 2nd Edition, London.

Weimann, Joachim (1994): Individual Behaviour in a Free Riding Experiment, in: *Journal of Public Economics* 54, S. 185-200.

Winter, Fabian / Rauhut, Heiko / Helbling, Dirk (2012): How Norms Can Generate Conflict: An Experiment on the Failure of Cooperative Micro-Motives on the Macro-Level, in: *Social Forces* 90, S. 919-948.

Wolf, Stephan / Dron, Cameron (2015): Intergenerational Sharing of Non-Renewable Resources: An Experimental Study Using Rawls's Veil of Ignorance, Freiburg. Online verfügbar unter <http://www.wipo.uni-freiburg.de/dateien/research/cen-papers/cen01-2015> [12.1.2018].

Wolf, Stephan / Lenger, Alexander (2014): Utilitarianism, the Difference Principle, or Else? An Experimental Analysis of the Impact of Social Immobility on the Democratic Election of Distributive Rules, in: Christoph Lütge / Hannes Rusch / Matthias Uhl (Hg.): *Experimental Ethics. Toward an Empirical Moral Philosophy*. Houndmills, S. 94-111.

Zimmermann, Ekkart (2015): Das Experiment in den Sozialwissenschaften. Entwicklungen und Chancen, in: Marc Keuschnigg / Tobias Wolbring (Hg.): *Experimente in den Sozialwissenschaften*. Soziale Welt, Sonderband 22. Baden-Baden, S. 17-33.

Leitfaden

I. Block: Experiment mit Produktionsspiel und Reform			
Leitfrage/Erzählaufforderung			
Erzählen Sie doch mal, wie hat Ihnen das Experiment gefallen?			
Inhaltliche Aspekte	Aufrechterhaltungsfragen	Konkrete Nachfragen	Hinweise/Instruktionen
Gesamtbewertung des Experiments Bewertung des Ergebnisses Bewertung des Experimentaufbau Bewertung der Produktionsbedingung Reform	Fällt Ihnen dazu sonst noch etwas ein? Gibt es sonst noch etwas? Und sonst? Fallen Ihnen noch weitere Punkte ein?	Waren Sie mit dem Experiment insgesamt zufrieden? Sind Sie mit Ihrem Abschneiden bzw. mit ihrer persönlichen Auszahlung zufrieden? Waren Sie mit den Spielregeln einverstanden, also mit der Möglichkeit die Einkommensverteilung in der zweiten Runde zu ändern? Wie fanden Sie die Produktionsspiele?	

II. Block: Experimentelle Ökonomik			
Leitfrage/Erzählaufforderung			
Was halten Sie persönlich von solchen ökonomischen Experimenten?			
Inhaltliche Aspekte	Aufrechterhaltungsfragen	Konkrete Nachfragen	Hinweise/Instruktionen
Teilnahmemotive Vorkenntnisse/Erfahrungen mit experimenteller Ökonomik Realitätsnähe von ökonomischen Experimenten Fachkulturen	Fällt Ihnen dazu sonst noch etwas ein? Gibt es sonst noch etwas? Und sonst? Fallen Ihnen noch weitere Punkte ein?	Was war Ihre Motivation an diesem Experiment teilzunehmen? Haben Sie zuvor schon einmal an einem solchen oder ähnlichen Experimenten teilgenommen? Verhält man sich anders, wenn man Experimentiererfahrung hat? Kennen Sie solche Experimente oder andere Experimente aus Ihrem Studium? Inwieweit spiegelt das Experiment die Realität wieder? War Ihnen beim Ablauf stets bewusst, dass es sich um ein Experiment handelte oder blendet man das irgendwie aus? Hätten Sie genauso abgestimmt, wenn es um mehr Geld gegangen wäre?	

III. Block: Gerechtigkeit und soziale Ungleichheit			
Leitfrage/Erzählaufforderung			
Fanden Sie persönlich das Experiment gerecht?			
Inhaltliche Aspekte	Aufrechterhaltungsfragen	Konkrete Nachfragen	Hinweise/ Instruktionen
Gerechtigkeitsmotive Soziale Ungleichheit Fachkultur Perspektivwechsel	Fällt Ihnen dazu sonst noch etwas ein? Gibt es sonst noch etwas? Und sonst? Fallen Ihnen noch weitere Punkte ein?	Sind Sie mit Ihrem persönlichen <i>Auszahlungsergebnis</i> zufrieden? Was wäre heute ein faires Einkommen gewesen? Was ist Ihrer Meinung nach die gerechteste Einkommensverteilung von den vier Auszahlungsmustern? Was ist generell eine gerechte gesellschaftliche Einkommensverteilung? Wie finden Sie es, dass sie im Produktionsspiel so gut abgeschnitten haben? Haben Sie damit gerechnet? Welches Fach studieren Sie? Wie hätten wohl Studierende aus anderen Fachbereichen abgestimmt? Wissen Sie, was für Fächer die anderen Versuchsteilnehmer studieren? Eine letzte Frage noch: Was denken Sie was wir herausfinden wollen?	Darauf achten, ob sie bereits im Block I beantwortet wurde! Möglichst ausführlich erzählen lassen Darauf achten, ob sie bereits im Block I beantwortet wurde!