Das Ultimatumspiel mit multiplen Respondern – Reproduktion eines Experiments aus Bern

An der Universität Leipzig,
Fakultät für Sozialwissenschaften und Philosophie,
Institut für Soziologie
im **Masterstudiengang Soziologie** eingereichte

Masterarbeit
zur Erlangung des akademischen Grades eines
Master of Arts (M.A.)

vorgelegt von

Jörg Neumann, B.A.
geboren am 30.08.1988 in Zwenkau
Matrikelnummer: 163410
Chopinstraße 10, 04103 Leipzig
obc_j-di@gmx.de

Betreuung durch Prof. Dr. Roger Berger
(Lehrstuhl Soziologie und Methodenlehre)

Danksagung

Ohne die Hilfe einiger Personen wäre die vorliegende Arbeit in dieser Form nicht möglich gewesen. Mein Dank gilt allen voran Professor Dr. Roger Berger für sehr wertvolle und motivierende Anmerkungen während der Betreuung der Arbeit. Nicht minder wichtig ist seine Unterstützung, indem er die notwendigen Ressourcen für die Durchführung der Experimente über das Leipziger Experimentallabor „LEx“ stiftete. Sehr dankbar bin ich auch Sonja Pointner für ein lehrreiches Seminar in Bern und ihre Hilfsbereitschaft und komplikationslose Bereitstellung der notwendigen Dateien für die Reproduktion. Außerdem danke ich Bastian Baumeister für Rat und Tat bei der Vorbereitung und Umsetzung des Experiments im Labor in Leipzig, vor allem mittels der Software z-tree und beim Erstellen eines adäquaten Stata-Datensatzes.

Für weitere nützliche Kommentare danke ich dem kritischen Auditorium meines Vortrags im Kolloquium „Aktuelle Themen und Probleme der empirischen Sozialforschung“ im Wintersemester 2013/2014 am Institut für Soziologie, insbesondere Professor Dr. Thomas Voss und Dr. Andreas Tutić.

Abstract Deutsch

Die vorliegende Arbeit reproduziert ein Experiment zum Ultimatumspiel mit multiplen Respondern, das zuerst an der Universtität Bern unter Leitung von Professor Dr. Axel Franzen und Dr. Sonja Pointner durchgeführt wurde. Ein künstlicher Proposer richtete dabei eine unfaire Aufteilung einer Geldsumme an die echten Probanden des Experiments. Ein Design widmete sich der Verantwortungsdiffusion bezüglich einer altruistischen Sanktionierung des Proposers unter gleichzeitig und unabhängig voneinander entscheidenden Respondern. Neben einigen Ungenauigkeiten und Fehlern bei der Umsetzung des Experiments liefern kontraintuitive Resultate triftige Gründe zur Wiederholung dessen. Allerdings deuten auch die Befunde der Reproduktion in die entgegengesetzte Richtung des „bystander effects“. Die Probanden zeigen erneut sehr häufig uneigennütziges Verhalten, das auf soziale Präferenzen hinweist. Die Daten suggerieren allerdings, dass es sich dabei – zumindest teilweise – um ein low-cost-Verhalten handeln kann.

Abstract English

Ultimatum game with multiple responders
– a reproduction of an experiment in Berne.

This thesis reproduced an experiment about an ultimatum game with multiple responders, first conducted at the University of Bern, Switzerland, under guidance of Professor Dr. Axel Franzen and Dr. Sonja Pointner. Thereby an artificial proposer offered an unfair split of money to the responding participants of the experiment. One design focused on the diffusion of responsibility with respect of an altruistic punishment towards the proposer and responders, who had to decide simultaneously and independently of each other. Besides some inaccuracies and mistakes while conductiing the experiment, counterintuitive results deliver reasonable arguments for a repitition. Nevertheless, also the reproduction’s findings point against the direction of the „bystander effect“. The subjects likewise show unselfish behaviour, indicating social preferences. However, the data suggest that this – at least to some extent – could be just low-cost-behaviour.

Inhaltsverzeichnis

Abstract Deutsch ... 3
Abstract English ... 3
Inhaltsverzeichnis ... 4
Abbildungsverzeichnis ... 6
Anhangsverzeichnis ... 6
Abkürzungsverzeichnis ... 6
1 Einleitung ... 7
2 Theorie ... 8
2.1 Soziale Präferenzen ... 8
2.1.1 Fairness ... 9
2.1.2 Reziprozität ... 11
2.1.2.1 Direkte Reziprozität ... 11
2.1.2.2 Indirekte Reziprozität ... 12
2.1.3 Altruismus ... 14
2.2 Sanctioning Dilemma und Verantwortungsdiffusion ... 17
2.2.1 Verantwortungsdiffusion ... 17
2.2.2 Exkurs: Asymmetrisches Freiwilligendilemma ... 21
3 Forschungsstand ... 22
3.1 Ultimatumspiel ... 22
3.2 Ultimatumspiel mit multiplen Respondern ... 24
3.2.1 Konkurrenz unter multiplen Respondern ... 25
3.2.2 Weitere Ultimatumspiele mit multiplen Respondern ... 26
4 Das Experiment in Bern ... 29
4.1 Forschungsdesign ... 29
4.1.1 Konkurrenzsituation ... 30
4.1.2 Verantwortungsdiffusion ... 31

4.2 Charakteristika des Designs ... 33
4.2.1 Das fingierte niedrige Angebot ... 33
4.2.2 Spielertypen ... 33
4.2.3 Die Spielstruktur ... 34
4.3 Durchführung ... 35
4.4 Ergebnisse ... 36
4.4.1 Univariat ... 36
4.4.2 Bivariat ... 37
4.4.3 Multivariat ... 39
4.5 Probleme des Experiments und Gründe für die Reproduktion ... 42
5 Hypothesen ... 45
6 Reproduktion ... 46
6.1 Modifikationen des Experiments gegenüber Bern ... 46
6.2 Durchführung ... 47
7 Ergebnisse ... 49
7.1.1 Univariat ... 49
7.1.2 Bivariat ... 49
7.1.3 Multivariat: ... 51
8 Methodendiskussion ... 56
9 Schlussbetrachtungen ... 58
Literaturverzeichnis ... 62
Anhang ... 67

Abbildungsverzeichnis

Abb. 1: Grundform des Freiwilligendilemmas nach Diekmann 1985 19
Abb. 2: Mini-Ultimatumspiel in extensiver Form. 23
Abb. 3: Ultimatumspiel mit multiplen Respondern in sequentieller Anordnung (extensive Form)... 31
Abb. 4: Ultimatumspiel mit multiplen Respondern in simultaner Anordnung (extensive Form)..... 32
Abb. 5: Ablehnungsverhalten der Responder in Bern. 40
Abb. 6: Ablehnungsverhalten der Responder in Leipzig. 54
Abb. 7: Vergleich des Ablehnungsverhaltens zwei simultan agierender Responder in Bern und Leipzig. 55

Anhangsverzeichnis

1) Deskriptive Übersicht unabhängiger Variablen der logistischen Regression (LEx), n=115 67
2) Deskriptive Übersicht unabhängiger Variablen der logistischen Regression (Bern) 67
3) Spieleranweisung 68
4) Fragebogen im Anschluss an das Experiment in Bern 72
5) Fragebogen im Anschluss an das Experiment in Leipzig 78

Abkürzungsverzeichnis

Abb.	Abbildung
d.h.	das heißt
DM	Deutsche Mark
ebd.	ebenda
etc.	et cetera, und so weiter
vgl.	Vergleiche
vs.	versus
z. B.	zum Beispiel

Anmerkung des Autors:

Werden Personenbezeichnungen aus Gründen der besseren Lesbarkeit lediglich in der männlichen oder weiblichen Form verwendet, so schließt dies das jeweils andere Geschlecht mit ein, sofern nicht anders vermerkt.

1 Einleitung

Die vorliegende Arbeit ist eine Reproduktion eines Experiments, das an der Universität Bern 2012/2013 im Rahmen eines Forschungsseminars unter der Leitung von Professor Dr. Axel Franzen und Dr. Sonja Pointner durchgeführt wurde. Konzeptioneller Ausgangspunkt ist dabei das aus der Spieltheorie stammende Ultimatumspiel. Mit dessen Hilfe ergründen Wissenschaftler verschiedener Disziplinen seit über drei Jahrzehnten das Verhalten der Menschen. Je nach Erkenntnisinteresse werden einzelne Elemente variiert und es ergeben sich zahlreiche Ausprägungen. Es zeigte sich recht häufig, dass die Akteure in Experimenten nicht nur ratioanl handeln, sondern auch soziale Präferenzen offenbaren, wenn auch teils aus Eigeninteresse motiviert. Ein Charakteristikum, dem bis dato noch nicht übermäßig Aufmerksamkeit geschenkt wurde, sind mehrere passive Spieler, sogenannte Responder. Im Vergleich mit dem „klassischen“ Ultimatumspiel stellt sich die Frage: Inwiefern wird das Verhalten des Proposers und der Responder durch die Präsenz mehrerer anderer Responder beeinflusst?

Die meisten Forschungsdesigns aus der Vergangenheit sahen vor, den Proposer mit einem echten Probanden zu versehen. Als ein Alleinstellungsmerkmal wird in der vorliegenden Arbeit zunächst darauf verzichtet, weil sonst die „Gefahr“ sozial erwünschter, fairer Angebote zu groß ist. Der natürliche Proposer wird durch einen Computer ersetzt, der sich nicht automatisch normativen Zwängen unterwirft. Ziel ist es, darüber ein unfaires Angebot zu erzielen, das eine potenzielle Ablehnung seitens der echten Responder provoziert. Dabei wird ihre genaue Zahl verändert und in Abhängigkeit davon die Reaktion beobachtet. Im Anschluss daran sind die Probanden selbst in der Situation, einen Geldbetrag unter sich und diesmal fiktiven Mitspielern aufzuteilen.

Die Arbeit grenzt sich also zu anderen Studien ab, die die ad-hoc-Einhaltung von Fairnessnormen in „vorauseilendem Gehorsam“ untersuchen, oder solchen, bei denen die Kostenbeteiligung an einem gemeinsam genutzten Sanktionssystem im Fokus steht. Hier ist insbesondere die individuelle, uneigennützige Wiederherstellung der normativen Ordnung und außerdem die indirekte negative Reziprozität von immanenter Bedeutung.

Im weiteren Verlauf der Arbeit werden zuerst die theoretischen Grundlagen der sozialen Präferenzen sowie dem Freiwilligendilemma dargelegt. Anschließend wird das einfache Ultimatumspiel und der Forschungsstand zu Experimenten mit multiplen Respondern behandelt. Danach folgt die Berner Versuchsanordnung inklusive der Ergebnisse, Probleme und der Begründung für eine Reproduktion. Die empirisch zu überprüfenden Hypothesen münden ihrerseits in den methodischen Teil der Arbeit. Es wird nochmals auf das in der Reproduktion realisierte Design eingegangen, ehe deren Resultate referiert werden. Das Ende bildet eine Diskussion der Ergebnisse.

2 Theorie

2.1 Soziale Präferenzen

Empirisch auftretende Abweichungen von verhaltensökonomischen Vorhersagen sind nicht unbedingt beschränkt rational oder gar irrational, also Antonym zum Egoismus, sie werden gemeinhin als Nachweis für soziale Präferenzen interpretiert (Pointner 2012, S. 69). Wenn man der Argumentation folgt, dass sie ebenso wie egoistische Präferenzen entscheidungstheoretische Axiome erfüllen, dann schließen im Sinne des Konsistenzkriteriums beide einander nicht aus, sondern sind miteinander theoretisch vereinbar (Diekmann 2014, S. 60). Entsprechende Modifikationen der standardökonomischen Modelle berücksichtigen dies, indem sie zusätzlich soziale Nutzenfunktionen formalisieren. Über die intraindividuelle Präferenzordnung sind dann eindeutige Entscheidungen erlaubt, wonach sich ein Akteur rational verhält, wenn er seinen sozialen Präferenzen folgt (Pointner 2012, S. 49).
Damit ist verbunden, dass beobachtetes faires oder altruistisches Verhalten oft aus strategischen, also eigentlich eigennützigen Gründen erfolgt. So kann die Befolgung einer Norm aus sozialer Erwünschtheit, Reputationsaufbau oder einem mental empfundenen „warm glow“ resultieren (Pointner 2012, S. 45). Für eine überzeugende Reputation benötigt es für gewöhnlich langfristige Investitionen, während ein „schlechter Ruf“ wesentlich schneller entsteht[1].

Soziale Präferenzen sind genau genommen Präferenzen für soziale Normen. Den meisten Definitionen sozialer Normen sind folgende Aspekte gemeinsam: 1) eine Verhaltensregelmäßigkeit, die 2) auf einer sozial geteilten Erwartung daran basiert, wie jemand sich verhalten sollte, und selbst Auslöser ist für 3) die Bekräftigung des vorgeschriebenen Verhaltens mittels informeller sozialer Sanktionen (Fehr und Gächter 2000, S. 166). Daher kann eine soziale Norm als seine Art öffentliches Gut in puncto Verhalten gedacht werden, zu dem jeder einen Teil beitragen sollte, indem er der Norm Folge leistet und Normbrüche auch unter eigenen Kosten ahndet. Soziale Normen werden natürlich nicht bedingungslos befolgt. “If norm adherence or norm enforcement becomes more costly, norms are more likely to break down” (Fehr und Fischbacher 2004, S. 189).
Drei Normen sind im Rahmen dieser Arbeit und auch sonst gesellschaftlich besonders bedeutsam, nämlich Fairness, Reziprozität und Altruismus. Die bloße Existenz einer sozialen Norm mit ihrem Zweck für die Gesellschaft stellt unterdessen ihre Befolgung noch nicht sicher, sondern es muss ein

[1] Eine Redewendung aus dem Volksmund verdeutlicht dies: „Wer einmal lügt, dem glaubt man nicht“.

Anreiz dafür bestehen (z. B. eine glaubwürdige Sanktionsandrohung). Gewisse externe[2] und interne[3] Auslösemechanismen sind aber in Experimenten entweder schwierig herzustellen oder oft nicht beobachtbar (Pointner 2012, S.39). Das Verhalten ähnelt sich oft und kann nicht immer streng voneinander getrennt werden. Obschon Fairness eine schwächere Präferenz als Reziprozität sein kann, dient es für gewöhnlich als Stellvertreter dafür (Fehr und Fischbacher 2002), denn es ist ein Teil von „fairness-driven behavior“ (Falk et al. 2003, S. 20). Wie in den nächsten Abschnitten ersichtlich wird, sind beide jedoch nicht identisch.

2.1.1 Fairness

Neben egoistischen Abwägungen beeinflussen auch Fairnesserwägungen die Entscheidungen der Akteure (Fehr und Gächter 2000, S. 163). „Geld mag die Welt regieren, aber Fairnessüberlegungen zwingen dazu, notfalls materielle Einbußen in Kauf zu nehmen“ (Vieth 2003, S. 347). Fairness kann als genetisch fixierte Präferenz, über Sozialisation internalisierte Präferenz oder als durch Sanktion gestützte soziale Norm aufgefasst werden (Fehr und Fischbacher 2003; Bicchieri und Xiao 2009, S. 192).

Es existiert eine substanzielle Vielfalt darin, was Menschen für Auffassungen über Fairness besitzen und welches Gewicht sie solchen Betrachtungen beimessen. Faktoren wie der Arbeitsmarkt oder die Bildung modifizieren die Fairnesspräferenzen, hauptsächlich hinsichtlich der konkreten Fairnessideale. Wesentlich weniger Einfluss wird hingegen auf die Bedeutung ausgeübt, die Fairness an sich für die Menschen besitzt – sie ist hoch und weithin gültig (Cappelen et al. 2010).

Die Norm der Fairness erstreckt sich auf das Verfahren in einer Sache, einer Situation oder eines Ergebnisses (Pointner 2012, S. 35). Letzteres wird als „equity“ bezeichnet, im Sinne von Auszahlung beziehungsweise gleicher Verteilung des Geldes und genau das ist Gegenstand der Arbeit. Spieltheoretisch ist die Fairnessnorm mit der Gleichgewichtsauswahl verknüpft, das heißt es entwickeln sich soziale Erwartungen mit Koordination auf genau ein Gleichgewicht (innerhalb mehrerer Runden; Vieth 2003, S. 350).

2 „Eine soziale Norm der Fairness […] kann also einer der Ausgangspunkte für die Entstehung einer Fairnesspräferenz und Auslöser für entsprechendes Verhalten sein“ (Pointner 2012, S. 44).

3 Neben den externen bestehen noch interne Steuerungsgründe für prosoziales Verhalten, z. B. Tugenden und Prinzipien (Zimbardo und Gerrig 2008, S. 682), die als handlungsleitende Maximen, die ihre übergeordneten Wurzeln in Moral, Ethik oder Religion besitzen, wie z. B. der kategorische Imperativ von Kant. Diese stehen hier aber nicht im Fokus.

Das Fairnessurteil, das Menschen über die Taten einschließlich ihrer Konsequenzen einer anderen Person ausbilden, beeinflusst wiederum die Präferenzen für (eigene) Taten und Allokationen (Camerer 2003, S. 114). Langfristig gesehen kann sich faires Verhalten auch aus rationalem Eigeninteresse begründen, z. B. um Fairness zu signalisieren und sozial erwünscht zu handeln – somit ist es sozial eingebettet (Vieth 2003, S. 351). Andererseits kann das Resultat einer Entscheidung ab und zu unfair sein, obschon die Intention des Mitspielers eigentlich fair war (Pointner 2012, S. 50).
Im Gegensatz zu Normen der Kooperation gehört Fairness zur Kategorie der disjunkten sozialen Normen, weil die Adressaten und Profiteure der Norm auseinanderfallen, so dass gemeinsame Vorteile zunächst nicht gegeben sind (Vieth 2003, S. 361). Eine ähnliche personelle Inkongruenz wird hinsichtlich des „sanctioning dilemma“ im späteren Abschnitt relevant.

Die empirischen Untersuchungen der Fairnessnorm im Labor mittels Ultimatumspielen und insbesondere Diktatorspielen[4] ergeben eine starke Kontextabhängigkeit der Fairness. Engel (2011) findet in einer Meta-Analyse Abgaben von durchschnittlich 28 % des aufzuteilenden Betrags, indes mit einer erheblichen Varianz. Faires Handeln tritt häufiger auf, wenn die Situation es auch nur in geringem Maße nahelegt, allerdings seltener bei anonymen Experimenten, während dennoch ein beachtlicher Teil an Personen Abgaben leistet. Die nicht vernachlässigbaren Abgaben in Diktatorspielen liegen zwar über den standardökonomischen Vorhersagen, aber unter denen der Ultimatumspiele. Fairness scheint sich also insgesamt (gegenüber Effizienz) behaupten zu können, tritt aber reduziert auf, wenn der Aufteilung eine Verdienstphase vorausgeht (Pointner 2012, S. 85ff.).[5]
Auch wenn die Designs keinen Anlass dazu geben, kann Fairness selbstverständlich in anonymen Experimenten mit einmaligen Interaktionen auftreten, trotz oder gerade weil Probanden versehentlich ihre üblichen Handlungsroutinen beziehungsweise Verhaltensregeln im Labor anwenden (Pointner 2012, S. 255).

[4] Das Diktatorspiel ist eine Variante des Ultimatumspiels, bei dem der Responder keine Möglichkeit zum Veto besitzt. Es ist daher von großer Bedeutung zur Untersuchung fairen Verhaltens.

[5] Genau das bleibt außen vor, da es sich hier bei den Geldbeträgen um „windfall gains“ handelt.

2.1.2 Reziprozität

„Humans are the champions of reciprocity“, so Nowak und Sigmund (2005, S. 1291). Die Thematik der Reziprozität spielt seit jeher eine wichtige Rolle in den Humanwissenschaften und ist in soziologischer Hinsicht vor allem Gegenstand klassischer Arbeiten von Simmel, Malinowski, Mauss und Gouldner (Diekmann 2004, S. 487). Reziprokes Verhalten ist mit fairem Handeln verwandt, jedoch nicht identisch: „If someone voluntarily shares a cake with somebody else, he or she adopts an approach of unconditional fairness. Reciprocity, on the other hand, is conditional fairness“ (ebd., S. 489). Der Reziprozität wohnt der Gedanke der Wechselseitigkeit inne (Rabin 1993, S. 1281), sie ist Teil des Ausgleichs zwischen Akteuren beruhend auf vergangenem Verhalten untereinander, das heißt der kompensatorischen Reziprozität oder „ausgleichenden Gerechtigkeit“ (Pointner 2012, S. 37).[6]

2.1.2.1 Direkte Reziprozität

Wenn Ego von Alter einen Gefallen erhält, befindet er sich zunächst in einem „Zustand psychischen Unwohlseins, bis [er] sich revanchieren [kann]“ (Zimbardo und Gerrig 2008, S. 681; Weber et al. 2004, S. 296; Fehr und Gächter 2000, S. 161). Er kann dies identisch zurückzahlen, nach dem Prinzip „tat-for-tat“ oder „you scratch my back, and I'll scratch yours“ (Nowak und Sigmund 2005, S. 1291). Alternativ kann er die Schuld durch ein gleichwertiges Äquivalent („tit-for-tat“) begleichen (Diekmann 2004, S. 489). Das Prinzip gilt sowohl positiv oder kooperativ im Sinne einer Gegenleistung, aber auch in negativer Hinsicht als Vergeltung gemäß „Auge um Auge, Zahn um Zahn“. Bezogen auf das dargestellte Experiment bezeichnet diese „starke Reziprozität“ die Tendenz von Ego, eigens erfahrenes kooperatives Verhalten zu belohnen und nicht-kooperatives zu bestrafen, sogenanntes second-party-punishment (Nowak und Sigmund 2005). Reziprozität ist eine evolutionär stabile Strategie, deren Schlüsseleigenschaft ihre Unabhängigkeit von zukünftigen Interaktionen ist (Gintis et al. 2003, S. 153f.). Daneben besteht die aufrechterhaltende Reziprozität („weak reciprocity“), bei der längerfristige Beziehungen zwischen Individuen eingegangen werden, wobei die Leistungen und Gegenleistungen nicht akkurat miteinander verrechnet werden, da es zum Teil gar nicht möglich ist (Pointner 2012, S. 37f.).

---

[6] Die Ungleichheitsaversion, also die Präferenz für Fairness ist die schwächer als die für Reziprozität (Fehr und Fischbacher 2002, S. 3). Die Verpflichtung durch reziprokes Verhalten scheint verbindlicher zu sein als die durch bedingungslose Fairness, was letztlich noch empirisch ergründet werden muss (Diekmann 2004, S. 502).

In seiner Grundform beinhaltet das Konzept, dass beide am Tausch beteiligten Parteien Rechte und Pflichten besitzen – es hat demnach eine große Relevanz für Rollensysteme. Aus den Wiederholungen der Transfers kann sich selbst eine generalisierte moralische Norm etablieren, die Verpflichtungen für erhaltene „Leistungen“ festlegt und dadurch die Stabilität des reziproken Austausches fördert (Gouldner 1984, S. 93ff.). Sie selbst ist zwar keine notwendige Ausgangsbedingung zur Herstellung reziproker Kooperation. Ein großer der Schatten der Zukunft kann allerdings rein egoistische und rationale Akteure eben dazu veranlassen (Diekmann 2004, S. 490). Die Norm der Reziprozität ist wahrscheinlich universell (aber nicht zwingend uneingeschränkt gültig) und kulturell bedeutsam, ganz ähnlich dem Inzest-Tabu (Gouldner 1984, S. 97). Über sie selbst und auch den Glauben an sie können egoistische Prinzipien einem sozialen Tauschsystem zur Stabilität verhelfen.

Weil rein ökonomische Ansätze Schwierigkeiten darin hatten beziehungsweise haben, die Art und Weise des initialen Austausches von Gütern zu erklären, dient die Reziprozitätsnorm als Auslösemechanismus (ebd.: 102).[7] Sie kann also strategisch motiviert sein, wenn Ego die Hilfe anderer für sich selbst erwartet, sollte Ego zunächst anderen helfen. Egoistischere Akteure werden sich daher ebenfalls in wiederholten dyadischen Beziehungen und größeren Gruppen durch (indirekte) Reziprozität eher kooperativ verhalten.

Es ist aber zu beachten, dass über das Konzept der direkten Reziprozität nicht jede Art der Kooperation bei sozialen Arten erklärt werden kann (Zimbardo und Gerrig 2008, S. 682), denn nicht immer ist eine Rückzahlung einer altruistischen Tat möglich. Man kann nicht erwarten, dass der Fremde, dem man das Leben gerettet hat, sich dafür revanchiert und gleichermaßen das Leben rettet. Zur Erklärung dessen verhilft ein weiteres Konzepts, das der indirekten Reziprozität.

2.1.2.2 Indirekte Reziprozität

Menschen können antizipieren, für ihre gegenwärtigen altruistischen Taten zukünftig selbst Nutznießer derartiger Handlungen zu werden. Solch eine an andere gerichtete Revanche heißt „third party altruism” oder indirekte Reziprozität (Nowak und Sigmund 2005, S. 1291) und ist somit über mindestens einen weiteren Akteur vermittelt:

[7] Ohne weitere Regelungen würden egoistische Akteure vielleicht bei einem potenziellen Tausch ihr Gut behalten, aber das andere besitzen wollen, so dass sie in einem „Du zuerst!“ verharren und der Handel endlos hinausgeschoben wird.

"You scratch my back and I'll scratch someone else's" beziehungsweise „I scratch your back and someone else will scratch mine" (ebd.).[8] An diesem Punkt wird eine Möglichkeit der Reputationsbildung und moralischer Beurteilung tangiert, da es auch Dritte einbezieht, die wiederum Egos Status diskreditieren können, falls dieser zwar Hilfe erhält, sie aber nicht weiterleitet (Nowak und Sigmund 2005, S. 1291). Das Konzept ist also zwischen direkter Reziprozität und Kollektivgütern angesiedelt, denn einerseits betrifft es zwei Spieler unmittelbar (Sender und Nutznießer), andererseits muss es in einer größeren Gruppe gespielt werden, bei der die Reputation wirksam wird (ebd.: 1296).

Es ist unstrittig, dass Reziprozität unter bestimmten Bedingungen empirisch in Erscheinung treten und Verhalten dominieren kann (Fehr und Gächter 2000, S. 159f.). Es herrscht aber Uneinigkeit darüber, worin genau die Quelle für reziprokes Verhalten liegt.

Oft genug treten positive oder negative, direkte oder indirekte Reziprozität in einmaligen Austauschsituationen auf, ohne dass dabei Verwandte Egos oder signifikante Andere miteinander interagieren und Statusbildung möglich sein könnte, was sich in einer Reihe von Experimenten niederschlägt (Diekmann 2004, S. 490). Reziproke Kooperation erweist sich als recht erfolgreiche, jedoch weniger stabile Strategie, sofern mehrere Akteure miteinander im Austausch stehen im Vergleich mit dyadischen Beziehungen in wiederholten Spielsituationen (ebd.). Bedeutsam ist zudem das beobachtbare reziproke Verhalten der Akteure bei niedrigen Vorleistungen, also direkte und indirekte negative Reziprozität. Das gilt aber nicht für direkte und indirekte positive Reziprozität, bei der hohe Vorleistungen nicht honoriert werden (Pointner 2012, S. 252). Ergebnisse aus Ultimatumspielen (Güth, Schmittberger und Schwarze 1982) und zum Bestrafungsverhalten (Fehr und Gächter 2002) verdeutlichen, dass Adressaten unfairen Verhaltens eigene Kosten zur Vergeltung nicht scheuen.

Andere Studien decken den Eigennutz einer nicht unwesentlichen Zahl der Akteure auf, wobei trotzdem die Mehrheit der Akteure reziprok handelt (Fehr und Gächter 2000, S. 162). Ein negativer, selbst schädigender Akt der Reziprozität kann anderen Personen einen Nutzengewinn aus der Revanche bereiten (Diekmann 2008). Damit tritt die dritte wichtige soziale Präferenz in den Vordergrund.

[8] Ein bekanntes Beispiel sind die Regelmäßigkeiten der Nahrungsteilung beim indigenen Stamme der Ache von Paraguay, die selbst bei geringer Größe der Beute mehr als die Hälfte der Nahrung miteinander teilen, wenngleich ein hoher sozialer Druck dazu besteht (Fehr und Fischbacher 2003, S. 789).

2.1.3 Altruismus

Unter Altruismus versteht sich ganz allgemein eine bedingungslose Freundlichkeit, der kein empfangener Altruismus vorausgegangen sein muss, wie das bei Reziprozität als gleichartige Antwort der Fall ist (Fehr und Gächter 2000; Fehr und Fischbacher 2002). In anonymen Experimenten mit „one-shot"- Situationen oder auch letzten Spielrunden werden etwaige Abgaben als rein altruistisch begründeter Akt interpretiert, weil keine Gegenleistung des Mitspielers mehr zu erwarten ist (hauptsächlich bei Diktatorspielen).
Natürlicher oder hedonistischer Altruismus ist ein bereitwilliger Dienst, für den keine Gegenleistung erwartet wird, bei dem man sich emotional mit dem Empfänger der Leistung verbunden fühlt und sogar noch einen gewissen Zusatznutzen erfährt (Pointner 2012, S. 35f.). Perfekter Altruismus ist die gesteigerte, selbstloseste Form mit bloßer Zufriedenheit durch die Tat, weswegen sie sich seltener beobachten lässt beziehungsweise noch eher in Familien auftritt („kin altruism"; ebd., S. 36). Daneben besteht noch der moralische Altruismus, der durch (höherrangige) Normen oder Prinzipien hervorgerufen wird, wobei soziale Anerkennung durch die Befolgung der Norm relevant ist oder sein kann (ebd.). Somit wird bei allen altruistischen Handlungen die Wohlfahrt der Nutznießer gesteigert, was zu Lasten von Ego geschieht, der den Nettoverlust zu tragen hat (Elster 2006, S. 185). Fraglich ist, ob der häufig bei altruistischen Handlungen empfundene „warm glow" messbar ist und ob dies nicht eher ein Nebenprodukt des altruistischen Tuns ist, sondern tatsächlich das Ziel der Handlung (Elster 2005, S. 208).

Auch altruistische Akte berühren die anderen beiden sozialen Präferenzen Fairness und Reziprozität. Der bisherigen Datenlage nach ist altruistisches Verhalten ein adaptiver Zug, also eine angemessene Reaktion von Ego auf eine Handlung von Alter (Fehr und Heinrich 2003). Altruistische Belohnung normkonformen Verhaltens beziehungsweise die altruistische Bestrafung einer Normverletzung sind entsprechend nahe der starken direkten Reziprozität angesiedelt (Fehr und Fischbacher 2003, S. 785). Die aus sozialen Dilemmata, sequentiellen- und Diktatorspielen gewonnenen Daten zeigen altruistische Handlungen vergleichsweise vieler Personen (Diekmann 2008). Es ist aber durchaus möglich, dass die besondere Verantwortung eines Diktators für seine eigene Auszahlung und die der anderen eine Art widerwilligen Altruismus widerspiegelt. Jener veranlasst Menschen dazu, sich zur Abgabe verpflichtet zu fühlen, wenn sie es sonst von sich aus nicht tun würden (Cryder und Loewenstein 2012, S. 443).

Das altruistic punishment

Normverletzungen befähigen die Betroffenen dazu, die Devianten mit Sanktionen zu belegen (Przepiorka und Diekmann 2013, S. 1). „More generally, non-selfishly motivated punishment constitutes a powerful device for the enforcement of social norms and human cooperation" (Fehr und Fischbacher 2004, S. 189). Akteure fühlen sich unwohl, wenn sie eine nicht geahndete Normverletzung beobachten und empfinden Erleichterung, wenn Gerechtigkeit hergestellt ist. Altruistische Bestrafung liefert für das ausübende Individuum eine gewisse Belohnung und ist damit Motiv und Anreiz dafür („Rache ist süß").

Die ausgeübte Sanktion ist mit mehr oder minder hoch ausfallenden Kosten verbunden, aber nicht immer altruistisch. Langfristig könnte sich die kostspielige Strafe eines Trittbrettfahrers sogar materiell auszahlen. Daher entschließen sich Akteure mitunter dazu, diese Kosten zu übernehmen, wenn kein anderer sie bereits übernimmt (Raihani und Bshary 2011).[9] Eine echt altruistische Bestrafung genügt folgendem Prinzip: Die Probanden treffen entweder einmalig oder in mehreren Spielrunden zumindest nicht identisch aufeinander, das heißt nicht dieselben Paarungen kommen zustande. Die Bestrafung von B für eine unfaire Handlung wird durch A ausgeübt, so dass C in einer späteren Spielrunde vom Erfolg dieser „Erziehungsmaßnahme" profitiert („educative punishment"). Das Individuum A ist daher zum altruistic punishment veranlasst, weil die erwartete (persönliche) Befriedigung aus der Handlung eben dazu motiviert (Elster 2005, S. 208). Materielle Motive sind damit ausgeschlossen, denn "[...] the act of punishment, although costly for the punisher, provides a benefit to other members of the population by inducing potential non-cooperators to increase their investments" (Fehr und Gächter 2002, S. 139). Durch diesen Aufwand der übernommenen Kosten ergibt sich: „Punishers suffer a fitness disadvantage […] compared with nonpunishing contributors" (Boyd et al. 2003, S. 3532), teilweise auch im Vergleich mit Trittbrettfahrern (Nikiforakis 2013).[10]

[9] Genau dann handeln die Akteure reziprok altruistisch. Starke Reziprozität hingegen stellt einen mächtigen Anreiz zur Kooperation dar, selbst in einmaligen Situationen und auch wenn Reputationsgewinne ausgeschlossen sind. Stark reziproke Akteure werden diejenigen bestrafen, die defektieren und solche belohnen, die kooperieren (Fehr und Fischbacher 2003, S. 788).

[10] Diese Bestrafungen unterliegen zwei häufig verletzten moralischen Begrenzungen: 1) Wer im Glashaus sitzt, sollte nicht mit Steinen werfen. 2) Man sollte eigens erfahrenen Ärger nicht an Dritten auslassen (Elster 2005, S. 206). Insbesondere letzteres betrifft also indirekte negative Reziprozität.

Nach intensiver Erforschung bestehen gegenwärtig vier relativ robuste Ergebnisse zum altruistic punishment (Nikiforakis 2013):

1) Selbst bei einmaligem Aufeinandertreffen sind Probanden bereit, auf eigene Kosten zur Kooperation oder gegen Trittbrettfahren beizutragen.
2) Altruistic punishment ist ein gewöhnliches und untergeordnetes Gut, also sensibel bezüglich seiner Kosten. Das heißt: Je höher der Preis, desto geringer die Nachfrage und je geringer die Ressourcen von Ego, desto geringer die Nachfrage nach dem Gut. Menschen scheinen eher eine Präferenz für die aufkeimenden Emotionen zu besitzen anstatt der Präferenz für kostenintensives punishment.
3) Die von altruistic punishment ausgehende Bedrohung hält die Kooperation in Laborexperimenten aufrecht bei einer ausreichend geringen „fee-to-fine ratio“, d. h. die Sanktionskosten in Relation zur Reduktion der Güterausstattung des Zielobjekts. Dieses Verhältnis liegt für gewöhnlich bei etwa einem Drittel und reflektiert die relativen Sanktionskosten, d. h. je geringer der Quotient, desto billiger ist die Bestrafung.
4) Die Gefahr einer Reaktion auf altruistic punishment (sogenanntes counter-punishment), ist eine wichtige Determinante für die Bereitwilligkeit zum Engagement in altruistic punishment.

Weiterhin verstehen die altruistischen „Richter” ganz eindeutig das Gewinnpotenzial, das im Falle einer möglichen Reputationsbildung aus dem Ruf als „harter Verhandlungspartner“ resultiert (Fehr und Fischbacher 2003, S. 787), sie kalkulieren also durchaus eigennützig. Trotzdem lassen sich sogar dann kostspielige Bestrafungen unfairer Spieler feststellen, wenn die Spielsituation nicht wiederholt und so innerhalb des Spiels keine Reputation aufgebaut werden konnte (Fehr und Gächter 2002, S. 139; Przepiorka und Diekmann 2013, S. 1). Das altruistic punishment tritt auch unter Anonymität auf, und zwar wegen emotionaler Gründe, aus Ärger über fehlende Kooperation anderer Mitspieler, als eine nicht-rationale Reaktion oder um faire Aufteilungen zu etablieren (Fehr und Gächter 2000; Fehr und Fischbacher 2003; 2004).

Es zeigte sich außerdem, dass das altruistic punishment nach Verletzung der Fairnessregel auch von „Zuschauern“ beziehungsweise dritten Parteien ausgeübt wurde (third-party-punishment in Verbindung mit indirekter Reziprozität), die nicht direkt davon betroffen sind und für die eine Durchsetzung ausschließlich Kosten verursacht. Je mehr die Vorgaben des Proposers unter die Fairnessmarke fielen, desto mehr wurde er von den Unparteiischen dafür bestraft – eine solche Sanktion wurde von den Respondern auch erwartet (Fehr und Fischbacher 2003, S. 786). Die Sanktion fällt durch den zweiten Spieler (Responder) zwar stärker aus als durch die unbeteiligte Drittpartei, den-

noch kommt letzteren eine wesentliche Rolle bei der Durchsetzung von Normen zu. Solch ein Verhalten lässt sich – wenn auch nicht so häufig – selbst in Feldstudien beobachten (Balafoutas und Nikiforakis 2012).
Empirisch führt das altruistic punishment in Gruppen also zu einer besseren Kooperation und der Wiederherstellung des normkonformen Zustandes, weil unfaire Spieler von den Anderen ausgeschlossen werden (Boyd et al. 2003, S. 3535). Allerdings: „[At] a cost to the punisher, and thus can be viewed as a social dilemma" (Nowak und Sigmund 2005, S. 1293; Dawes 1980; Barclay und Kiyonari 2014).

2.2 Sanctioning Dilemma und Verantwortungsdiffusion

Soziale Dilemmata sind allgemein durch zwei Merkmale gekennzeichnet: „(a) the social payoff to each individual for defecting behavior is higher than the payoff for cooperative behavior, regardless of what the other society members do" (Dawes 1980, S. 170) und (b) alle Teilnehmer der Gemeinschaft erzielen einen geringeren Nutzen, wenn alle defektieren im Vergleich zur Kooperation aller Beteiligten (ebd.). Die Akteure müssen zunächst einmal die Natur der Dilemmasituation erkennen, damit sowohl materielle Anreize als auch moralische, normative Belange Berücksichtigung finden. Wichtiges Merkmal ist die Unwissenheit über die Kooperation der anderen Gruppenmitglieder, demzufolge müssen die Akteure individuell eine genaue Vorstellung darüber entwickeln (Dawes 1980, S. 170; Nikiforakis 2013). In sozialen Situationen dieser Art sind es materielle Anreize, die Menschen zur Defektion verleitet, während es andere Nutzwerte in Verbindung mit sozialen Normen und Altruismus sind, die zu Kooperation anregen (Dawes 1980, S. 190). Einer der internen Auslöser prosozialen Verhaltens in solchen Situationen ist die soziale Verantwortlichkeit, die sich kurz formulieren lässt: Ego leistet Hilfestellung für denjenigen, der sie benötigt.

2.2.1 Verantwortungsdiffusion

Falls Passanten Zeugen eines Unfalls oder Überfalls werden, hoffen die meisten oder gar alle davon, dass sich jemand findet, der die Kosten (K) auf sich nimmt und dem Opfer hilft. Angestoßen wurde die Erforschung dieser Problematik durch den Mord an Kitty Genovese, der medial viel Aufmerksamkeit erzeugte und erstmals von Darley und Latané (1986) aufgegriffen wurde. Sie formulierten den Begriff „bystander effect", wonach mit steigender Zahl der Passanten die Wahrscheinlichkeit des Engagements sinkt.

Bereits ab einer Gruppengröße von drei Personen (ein „Opfer“, zwei Passanten) tritt dieser Effekt in Erscheinung und die Wahrscheinlichkeit für Hilfe sinkt auf 0.82 und liegt bei fünf Personen nur noch bei 0.62, Tendenz weiter fallend.

Die Beobachter der Situation sind erleichtert, wenn sich mindestens ein anderer Passant findet, um Hilfe zu leisten. In diesem Moment wird ein entsprechendes Kollektivgut produziert und die unbeteiligten Zuschauer erhalten einen Nutzen (U) – neben ihrem Anteil am öffentlichen Gut ebenfalls einen “warm glow” oder auch “psychic gain of U”, also ein beruhigtes Gewissen (Diekmann 1985, S. 606). Jedes Individuum spekuliert auf das Engagement der zahlreichen anderen Akteure. Wenn sich in der gesamten Gruppe niemand findet, einen Beitrag zu leisten, dann kommt das gemeinschaftliche Gut nicht zustande und alle Personen werden benachteiligt, der Schaden entsteht (ebd.). Das Dilemma besteht darin, ohne weitere Absprache einen engagierten Freiwilligen zu finden, daher auch der Name dieses Problems: „Freiwilligendilemma“.

Kein rationaler Akteur besitzt dabei einen Anreiz, die Strategie Kooperation (Zivilcourage zeigen, eine Sanktion ausüben etc.) zu spielen, da sie eine Differenz aus Kosten und Nutzen beinhaltet. Die Taktik Defektion („bystander“ bleiben) besitzt diesen Kostenfaktor nicht und ist auch keine dominante Strategie. Sie ist nicht die beste Antwort und garantiert für alle Strategien der Mitspieler eine mindestens genauso große Auszahlung. Dennoch ist sie rational unter der Annahme eines Spielers, dass andere Akteure die MaxiMin-Strategie[11] spielen (Diekmann 1985, S. 606f.). Es existiert in dieser Konstellation keine dominante Strategie, wonach also kein Akteur ein Interesse an einer Devianz davon hätte, unabhängig von den Spielzügen der anderen. Aber es bestehen N asymmetrische reine Gleichgewichte mit pareto-optimalen Auszahlungen, bei dem ein Spieler der Freiwillige ist, während der Rest defektiert (Diekmann 2009).

Im symmetrischen Fall sind die Kosten (K) und der Nutzen (U) aller Strategien für alle Akteure (N) gleich, das heißt sie empfangen dieselbe Anfangsausstattung und haben denselben Nutzen vom kollektiven, öffentlichen Gut. Daraus ergibt sich eine Fairnessregel, nach der alle dasselbe beitragen und verdienen (Nikiforakis 2013 in Anlehnung an Fehr und Gächter 2002).

[11] Maximierung der eigenen Mindestauszahlung in Abhängigkeit der Strategien der anderen Akteure (Camerer 2003, S. 25ff.; Diekmann 2009, S. 234). Heißt: wenn alle kooperativ handeln, ist Defektion individuell reizvoll.

Spielerzahl / Strategien	0	1	2	…	N-1
Kooperation	U-K	U-K	U-K	…	U-K
Defektion	0	U	U	…	U

Abb. 1: Grundform des Freiwilligendilemmas nach Diekmann 1985 mit U > U – K > 0; für mindestens 2 Spieler.

Mittels dieses spieltheoretischen Modells lassen sich die Wahrscheinlichkeiten für beide Strategien herleiten (Diekmann 1985). Im Falle von Defektion ist dies $q = \sqrt[N-1]{K/U}$ und für die Kooperation gilt $p = 1 - \sqrt[N-1]{K/U}$ (gemischtes Nash-Gleichgewicht). Daraus geht hervor – genau wie bei Darley und Latané (1968) – dass die Wahrscheinlichkeit zur Kooperation individuell betrachtet eine monoton fallende Funktion von Gruppengröße und Kosten-Nutzen-Verhältnis ist, woraus sich später Hypothesen ableiten lassen.

Das komplexe Koordinationsproblem besteht eben darin, dass das Gleichgewicht (ein Akteur erklärt sich freiwillig bereit, die anderen defektieren) bei symmetrischen Ausgangssituationen ohne Absprache und verbindlichen Vertrag nicht erreichbar ist (Diekmann 2009, S. 125). Eine offizielle Institution könnte die im Dilemma vorherrschende Unsicherheit unter den Probanden ausräumen. Eine einklagbare Vertragslösung wäre also immer besser als individuelles rationales Handeln (ebd.). In dieselbe Richtung dringen materielle Anreize oder Belohnungen für den Freiwilligen vor, was auch effizienter ist als unnötigerweise alle „bystander“ oder Trittbrettfahrer mittels Sanktion zur Kooperation zu zwingen (Barclay und Kiyonari 2014).

Im hier vorgenommenen Experiment ist jedoch nicht das Freiwilligendilemma um die Zivilcourage als Passant von Bedeutung. Die primäre Norm schreibt zunächst faires Verhalten des Proposers vor, wird aber durch das fingierte Angebot systematisch verletzt. Danach besteht für die Responder mit der Ablehnung des Angebots eine Sanktionsmöglichkeit, obschon sie diese nicht ergreifen müssen und auch anderen überlassen könnten. Somit existiert ein „sanctioning dilemma of N actors facing a norm violation“ (Diekmann 1993, S. 76). Die Strategie der Kooperation entspricht dem mit Kosten verbundenen „altruistic punishment“, also der Ablehnung des niedrigen fingierten Angebots[12], welches zugleich die Lösung des sanctioning dilemma darstellt (Diekmann 2004, S. 503).

[12] „Gossiping“ (zu Deutsch „Klatsch und Tratsch“) ist eine günstige, aber effektive Variante der Bestrafung von Trittbrettfahrern (Egas und Riedl 2008), obschon durch das vorliegende Design ohnehin unterdrückt.

Die für diese Arbeit wichtige Frage zum experimentellen Design ist also: Welcher der multiplen Responder übernimmt die Kosten des altruistic punishment des Proposers zur Wiederherstellung der sozialen Norm, von der alle anderen Responder profitieren?
Prinzipiell reicht bereits ein Akteur zur Vergeltung aus. Es müssen nicht alle geschädigten Gruppenmitglieder eine Sanktion ausführen und somit ihre eigenen Auszahlungen minimieren, weil damit auch kein zusätzlicher Nutzen generiert wird (Diekmann 1985). Der Schaden entsteht, wenn sich alle auf die anderen verlassen, so dass der Proposer für sein unfaires Angebot nicht gerügt wird (Przepiorka und Diekmann 2013, S. 1).
Die freiwillige, altruistische Bestrafung rührt aus den Überlegungen zu dem Verhältnis von Kosten und Nutzen (Wirkungsgrad) und ebenso aus emotionaler Erregung (Egas und Riedl 2008). Es wird insbesondere dann aufrechterhalten, je günstiger die Umstände sind, d. h. geringe Kosten für den Ausübenden und zugleich ein hoher Wirkungsgrad bei dem Bestraften (ebd.). Nichtsdestotrotz verursacht die Bestrafung abweichenden Verhaltens beiden Beteiligten monetäre Kosten und die Sanktionsdrohungen sind daher auf materieller Ebene eigentlich unglaubwürdig (Voss 2001). Dennoch sind die Kosten für den Bestrafenden hier relativ gering, denn der entgangene Anteil ist wegen des unfairen Angebots recht klein und trotzdem ist die Wirkung beim Proposer relativ stark, weil dieser seinen überproportionalen Anteil verliert.

Etliche empirische Studien verschiedener Disziplinen befassten sich mit dem Effekt der Verantwortungsdiffusion, neben der Sozialpsychologie auch die experimentelle Ökonomie, die Organisationsforschung und sogar in „computer mediated discourse“ (Barron und Yechiam 2002, S. 509). Zahlreiche kontrollierte, strukturell modifizierte Experimente erhärten und bestätigen fast durchweg den Effekt der Verantwortungsdiffusion. Trotzdem verhalten sich die Probanden häufiger kooperativ, als es das gemischte Nash-Gleichgewicht vorgibt (Diekmann 2009, S. 127).
Oft wird der Effekt der Verantwortungsdiffusion als Versagen menschlichen Mitgefühls angesehen. Die Ergebnisse legen hinsichtlich des moralischen Spielraums nahe, dass Personen sich häufig nicht um die Auszahlungen anderer kümmern, sondern vielmehr darum, selbst nicht eigennützig zu sein (Cryder und Loewenstein 2012, S. 443). Interne Impulse zu prosozialem Handeln werden oft genug gedämpft, weil die Verantwortung an eine Autorität abgetreten werden kann. Nichtsdestotrotz besteht die Tendenz mancher Personen, dennoch unliebsame Aufgaben zu übernehmen: „One may have a sense of ‚duty‘ which acts as a motivation.“ (Charness 2000, S. 376).

2.2.2 Exkurs: Asymmetrisches Freiwilligendilemma

Der empfangene Nutzen aus der Hilfeleistung in der Notsituation kann von Kultur zu Kultur variieren. „In einer egoistischen Gesellschaft wird der Wert des Kollektivguts geringer ausfallen als in einer solidarischen Gesellschaft“ (Diekmann 2014, S. 53). In derlei asymmetrischen Freiwilligendilemmata schwanken ganz individuell das Interesse am kollektiven Gut oder auch die Produktionskosten mit $U_i > (U_i – K_i) > 0$. Der stärkste Spieler erhält einen individuell höheren Nutzen oder hat geringere Kosten und ist damit besser als der Rest: $U_k – K_k > U – K$. Es wäre im Falle jenes Dilemmas nur allzu naheliegend, wenn die Beteiligten einwilligen, das „stärkste“ Gruppenmitglied auszubeuten (d.h. die Bestrafung ausüben zu lassen) und damit die Regel(n) am effektivsten und effizientesten zu untermauern (Przepiorka und Diekmann 2013).

Für das asymmetrische Experiment zeigen die Daten, dass mit höherer Differenz der Auszahlungen zwischen starken und schwachen Spielern die stärkeren Spieler wahrscheinlicher und schwächere weniger wahrscheinlich Kooperation auswählen (ebd.: 84). Falls die Kosten den Nutzen übersteigen ($K > U$), dann gleicht das Dilemma eher dem „Missing Hero Dilemma“ oder auch Schellings „Matratzenproblem“[13]. Weil dabei die Differenz aus Nutzen und Kosten für alle Beteiligten kleiner als Null ist ($U – K < 0$), ist Defektion eine dominante Strategie und die „Helden“ sind rar, wenn sie gerade benötigt werden (Diekmann 2009, S. 131). Diese strukturelle Variante könnte im vorliegenden Experiment gegeben sein. Jedoch ist das nicht allein aufgrund der Auszahlungen ersichtlich, da der individuelle Nutzen („warm glow“ oder „psychic gain of U“) nicht messbar ist beziehungsweise auch nicht abgefragt wird.

[13] Arnold von Winkelried hat der Sage nach den Eidgenossen in einer Schlacht gegen die Habsburger im Jahre 1386 zum Siege verholfen, indem er sich heldenhaft aufopferte und sich selbst mit einem Bündel gegnerischer Lanzen aufspießte. Beim Beispiel von Schellings Matratzenproblem ist die Frage, wer mit seinem PKW anhält und eine herrenlose Matratze von der Fahrspur einer stark befahrenen Straße beiseite räumt (für letzteres siehe Diekmann 2009).

3 Forschungsstand

In diesem Kapitel der Arbeit werden die bisherigen empirisch gewonnen Erkenntnisse aus Experimenten zu Ultimatumspielen, insbesondere solchen mit multiplen Respondern dargelegt, ohne Anspruch auf Vollständigkeit.

3.1 Ultimatumspiel

Das Ultimatumspiel wurde ursprünglich von Güth konzipiert und erstmals 1982 experimentell überprüft (Güth et al. 1982). Seither entwickelte es sich zu einem „standard work horse in social science research“ (Güth und Kocher 2014, S. 11). Bei dieser strategischen Entscheidungssituation erhält ein erster Spieler, der „Proposer“, einen Geldbetrag oder ein anderes Gut, das er zwischen sich und einem anderen anonymen Spieler aufteilen soll. Letzterer, der „Responder“, ist abhängig vom getätigten Angebot und kann es lediglich annehmen oder ablehnen („take-it-or-leave-it”; ebd., S. 3). In der einfachsten Fassung des Spiels besteht für die Akteure nicht die Möglichkeit, miteinander zu kommunizieren oder über Verträge gewisse Vereinbarungen treffen zu können. Es kann nicht gefeilscht oder auf eine zweite Runde verschoben werden – eben ein Ultimatum. Wird die Offerte seitens Proposer vom Responder geduldet, so wird sie auch realisiert. Schlägt der Responder sie allerdings aus, so erhalten beide nichts und das Spiel ist beendet. In diesem Moment „[…] they essentially ‚burn money‘ […]“ (Güth und Kocher 2014, S. 2). Die Bestrafung durch die Ablehnung des offerierten kleinen Anteils kostet den Responder weniger als den Proposer (Nowak et al. 2000, S. 1773) und kann als altruistischer Akt gesehen werden, denn die meisten Menschen verstehen die Gleichaufteilung als (erstrebenswertes) faires Ergebnis (Fehr und Fischbacher 2003, S. 785).

Die auf Seite 23 abgebildete Graphik (Abb. 2) ist eine Variante des Ultimatumspiels, bei dem der Proposer nur zwischen fairem und aggressivem Angebot wählen kann (Camerer 2003, S. 31). Es heißt daher Mini-Ultimatumspiel, die Kontrollgruppe im Berner Experiment orientierte sich daran.

Die Konzeption der ursprünglichen Rational Choice liefert die Vorhersagen, wonach die Entscheidungsträger in jeder Situation ihren finanziellen Nutzen maximieren (Güth und Kocher 2014). Das bedeutet, der Proposer erteilt das kleinstmögliche Angebot, weil der Responder jedes beliebige Angebot annehmen wird, welches mindestens den kleinstmöglichen Teil ausmacht (Vieth 2003, S. 350). Es gibt drei Gleichgewichte, nämlich die Strategienprofile (U, (J, j)), (U, (J, n)) und (F, (N, j)), die wechselseitig beste Antworten sind und bei denen kein Spieler einen Anreiz zum Abweichen besitzt. „The obvious solution concept for ultimatum bargaining games is the subgame perfect equilibrium point” (Güth et al. 1982, S. 368).

Von den Gleichgewichten ist nur (U, (J, j)) teilspielperfekt, da es die Gleichgewichtseigenschaft sowohl für beide (auch die nicht-realisierten) Teilspiele – beginnend nach den Zügen U oder F des Proposers – als auch für das Gesamtspiel erfüllt. Die beste Antwort des Responders auf das unfaire Angebot (8/2) des Proposers ist immer noch die Annahme desselben, so dass ein Nash-Gleichgewicht, das zugleich teilspielperfekt ist, realisiert wird (Camerer 2003, S. 31f.).

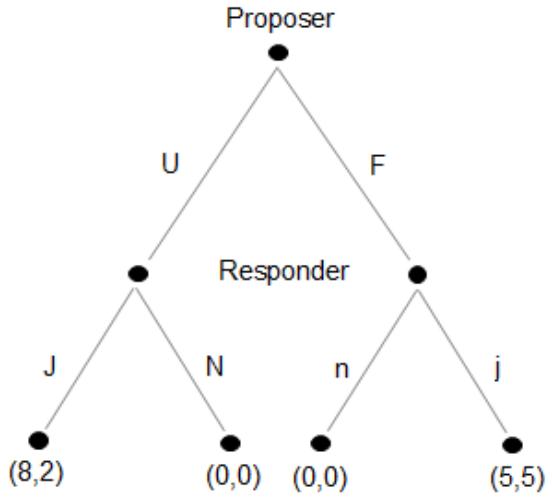

Abb. 2: Mini-Ultimatumspiel in extensiver Form.

Seit den ersten Untersuchungen zeigen die Ergebnisse unabhängig von der Spielform die Betonung individuell gerechtfertigter oder fairer Resultate seitens der Spieler. „Furthermore, the ultimatum aspect cannot be completely exploited since subjects do not hesitate to punish if their opponent asks for 'too much'" (Güth et al. 1982, S. 384). Die Proposer sind sich ganz offensichtlich des Ablehnungsrisikos bewusst und von eigenen Fairnessüberlegungen bestimmt, wenn sie Angebote erteilen (Güth und Kocher 2014, S. 3). Das zentrale und achtbar robuste Phänomen sind die Angebote von nahezu gleichen Aufteilungen (40-50 %), die fast immer angenommen werden. Die Akzeptanzraten in der Standardversion sinken mit kleiner werdenden Angeboten und erreichen zügig den Wert 0 für Anteile unter 20 % (ebd.). Es gibt fast keine Angebote von unter 10 % oder von hyperfairen 51-100 % (Camerer 2003, S. 49). Denn die Responder kümmern sich nicht nur um ihre eigene Auszahlung, sondern vergleichen diese auch mit der des Proposers und sind schnell frustriert, wenn diese wesentlich geringer ausfällt. Selbst bei der Untersuchung indigener Völker in 15 kleinen Stammesgesellschaften manifestieren sich diese Verletzungen der Vorhersagen des klassischen ökonomischen Modells (Henrich et al. 2001).

Eine kritische Würdigung des „klassischen" Ultimatumspiels: Es ist zwar nicht maximal realitätsnah, da die Akteure im Alltag selten anonym sind, der aufzuteilende Betrag weder vom Himmel fällt, noch

verschwindet und selten über mehrere „Runden“ verhandelt werden kann (Fehr et al. 2002, S. 54). Aber es liefert grundlegende Einsichten in Entscheidungsprozesse der Menschen und dient für andere Untersuchungen als Richtgröße beziehungsweise als Kontrollgruppe. Als Mittler zwischen Realität und Modell hat es zunächst die Vorhersagen der ökonomischen Modelle mit den empirischen Tatsachen konfrontiert. Danach versuchte man, die beobachteten Diskrepanzen zwischen Theorie und Empirie zu erklären, was – wie oben bereits angeklungen – zur Entwicklung neuer Theorien über sozialen Präferenzen führte (Güth und Kocher 2014, S. 10).

3.2 Ultimatumspiel mit multiplen Respondern

Ein gewichtiger Faktor und damit ein Argument für die Erweiterung des einfachen Ultimatumspiels um mindestens eine Person sind die sozialen Normen und Fairnessprinzipien, die sich auf mehr Beziehungsformen als eben nur Dyaden erstrecken. Denn wenn die Personen beim Fällen ihrer Entscheidung sich in Interdependenzen mit anderen Personen befinden, werden sie vermutlich andere Entscheidungen fällen als einzeln (Pointner 2012, S. 155). Die einfachste Erweiterung auf zwei Responder könnte als realitätsnahe Marktsituation zwischen einem Verkäufer und zwei potenziellen Käufern aufgefasst werden, wobei die Stärke des Proposers natürlich nicht von der Zahl der Responder abhängt (Fehr und Fischbacher 2002, S. 8). Dann ergeben sich unter anderem die Fragen, welche sozialen Motive und Normen fairer Aufteilungen gelten und ob die Spieler gleichermaßen zur Sanktion bereit sind, wenn eine unschuldige Partei davon betroffen ist (Camerer 2003, S. 80). Die meisten bisherigen Experimente ignorierten dies oft. Es bestehen aber dem Ultimatumspiel ähnliche Ausnahmen, die nachfolgend dargelegt werden. Es unterteilen sich zwei hier bedeutsame Versionen des Ultimatumspiels mit multiplen Respondern: einerseits mit sequentiell angeordneten Respondern und andererseits mit simultan entscheidenden Respondern[14].

[14] Die hier dargelegten Studien und Ergebnissse berücksichtigen die möglichen Veränderungen durch weitere Mitspieler, verfolgen aber teilweise andere Fragestellungen und haben daher modifizierte Versuchsanordnungen. Dennoch sind sie von Bedeutung, auch die derer mit Konkurrenz, obschon die Reproduktion sich nur auf das simultane Design konzentriert.

3.2.1 Konkurrenz unter multiplen Respondern

Die meisten Theorien sagen angesichts einer wie auch immer gearteten „Wettbewerbssituation“ einen Anstieg egoistischen Verhaltens vorher (Camerer 2003, S. 81). Vielleicht kümmern sich die Probanden um Fairness, aber sie handeln oft eigennützig, weil sie sich nicht sicher sein können, ob die anderen auch fair sind und weil sie Unfairness meist nicht ohne weiteres bestrafen können (ebd., S. 115). Die spieltheoretischen Vorhersagen der teilspielperfekten Gleichgewichte unterscheiden sich im Prinzip nicht vom einfachen Ultimatumspiel, der Proposer bietet den geringstmöglichen Betrag an und alle Responder weisen nichts zurück (Fischbacher et al. 2009, S. 503f.).

Grosskopf (2003) hat dies empirisch überprüft und die Spieler in Konkurrenz zueinander antreten lassen. Falls alle drei Spieler die offerierte Aufteilung von 100 Punkten ablehnen, erhält keiner der vier Spieler eine Auszahlung. Die Proposer erhielten keinerlei Informationen über Akzeptanzraten der Responder. Zum einen unterscheidet sich das Verhalten in den ersten Runden nicht signifikant, zum anderen steigen die geforderten Anteile im modifizierten Ultimatumspiel konstant und sind dazu noch höher als im traditionellen Ultimatumspiel, in dem sie auch noch häufiger abgelehnt werden. Tatsächlich führt einfache Konkurrenz dazu, dass sich die geforderten Anteile dem teilspielperfekten Nash-Gleichgewicht annähern. Ein möglicher Grund dessen ist die Unfähigkeit fairer Responder im Angesicht des Wettbewerbs, die Sanktion egoistischer Proposer durch das Ablehnen unfairer Angebote sicherzustellen, weil die konkurrierenden Responder eben solche womöglich annehmen (Fischbacher et al. 2009, S. 531).

Die Variation bestand bei Fischbacher et al. (2009) darin, dass zwei beziehungsweise fünf Responder simultan und in Konkurrenz zueinander ihre Entscheidungen fällen mussten. Sofern alle ablehnten, gingen alle leer aus. Sobald aber einer akzeptierte, dann erhielt er diese Auszahlung und der Proposer den übrig bleibenden Anteil. Wenn nur ein Responder ablehnte, dann empfing nur dieser nichts und von den anderen Respondern wurde per Zufallsauswahl die Auszahlung zugewiesen, der Proposer erhielt den verbleibenden Rest. Die Angebote fallen im Vergleich zum normalen Ultimatumspiel bei zwei Respondern um 29.1 % und bei fünf um 38.8 % geringer aus. Erwartungsgemäß sind die mittleren akzeptierten Angebote im Ultimatumspiel am höchsten (42.7 %), statistisch signifikant geringer bei zwei Respondern (25.5 %) und noch tiefer mit fünf Respondern (16.2 %). Es zeigte sich: Je geringer das Angebot, desto höher das Zurückweisungsrisiko des Proposers und je mehr Responder beteiligt sind, desto geringer ist selbiges Risiko. Dieser Rückgang fußt ausschließlich auf der Antizipation der Responder über das Verhalten der anderen Responder. Wenn Ego also annimmt, dass alle anderen Responder ebenfalls ablehnen, dann liegen die Ablehnungsraten im Konkurrenzfall genauso hoch wie im bilateralen Fall. Offenbar sind Fairnesspräferenzen und „Entscheidungsfehler“ enorm wichtig zum Verständnis, dass die Personen auch unter „Konkurrenz“ trotzdem von perfekter Rationalität abweichen.

3.2.2 Weitere Ultimatumspiele mit multiplen Respondern

Im Folgenden sind andere wichtige Studien aufgeführt, die mehr oder minder starke Ähnlichkeiten mit dem Design des Ultimatumspiels der vorliegenden Arbeit aufweisen.

Eine solche haben Oppewal und Tougareva (1992) vorgenommen, bei dem die Probanden aus Ost- beziehungsweise Westeuropa einen unterschiedlichen Bedarf der Lebenshaltungskosten aufwiesen. Wenn der erste Responder das Angebot akzeptiert, erhält der zweite Responder einen festen Anteil davon. Die Proposer boten wegen des dritten Spielers etwa 20 % mehr als im normalen Ultimatumspiel an. Sie zeigten also soziale Präferenzen und kümmerten sich um den weiteren Spieler, der keinen Einfluss auf die Aufteiungen hatte.

Bei Knez und Camerer (1995) tätigte der Proposer simultan und unabhängig voneinander eine Aufteilung von je 10°US-Dollar an zwei Responder (zwei separate Ultimatumspiele), mit Ausweichoptionen von 2 beziehungsweise 4°US-Dollar. Die Spieler waren in der Erhebungssituation nicht anonym, damit dadurch der Effekt des sozialen Vergleichs untereinander „besser“ zur Geltung kommen könne. Etwa die Hälfte der Responder hatte sich mit dem anderen Responder verglichen, was die Proposer nicht erwarteten beziehungsweise sich demgegenüber gleichgültig verhielten. Die zwei möglichen „Hintertürchen“ machten die Leute egoistischer in der Interpretation fairer Angebote, was häufiger zu Unmut führte. Die Ablehnungsraten lagen ähnlich wie in sonstigen Studien über 5 Runden beständig bei etwa 50 %. Die Responder verlangten mehr, sofern sie wussten, dass dem anderen Responder mehr angeboten wird, verglichen sich demnach untereinander. Die angegebenen minimal akzeptablen Angebote lagen allesamt über den Offerten der Proposer.

Güth und van Damme (1998) ließen den Proposer (X) den Geldbetrag (12 DM) zwischen sich, einem echten Responder (Y) und einem machtlosen Dummy-Responder (Z) aufteilen. Falls Y das Angebot annahm, erhielten alle drei die Auszahlung, ansonsten nicht. Das Desgin vereint die Aspekte des Ultimatumspiels (zwischen X und Y) und des Diktatorspiels (zwischen X und Z). Die Y-Spieler mussten benachteiligende Ungleichheiten zwischen sich und X einerseits und vorteilhafte Ungleichheiten zwischen sich und Z andererseits abwägen. Die rationale Annahme ist, dass echte Responder sich nicht um die übrige Aufteilung zwischen X und Z kümmern, was den Resultaten nach auch so ist. Die Proposer hatten keine starke intrinsische Motivation zur Fairness und keine einzige Ablehnung eines Angebotes seitens Y konnte eindeutig auf einen geringen Anteil für Z zurückgeführt werden, obschon sie bei nicht-monetären Anreizen geleitet waren. Selbige Studie wurde als Feldexperiment durchgeführt (Güth et al. 2007), in dem Leser von „Die Zeit“ teilnahmen. Die Gleichaufteilung war häufigstes Angebot und das mit höchster Akzeptanz. Offenbar war dem Responder Y die eigene Position wichtig, d. h. relativ niedrige Angebote wurden zurückgewiesen. Dennoch war die Reaktion von Y bedingt durch das Angebot an Z. Das Zeitungsexperiment enthüllte also im Gegensatz zu seinem Vorläufer im Labor ein Interesse gegenüber dem Dummyspieler.

Einen ähnlichen Versuch unternahmen Schmitt et al. (2008) an einer Militärakademie in den USA, wobei Z nur dann 5 von 15°US-Dollar erhielt, wenn die Aufteilung zuvor von Y angenommen wurde, ansonsten erhalten alle nichts. Die Vermutung war: „[T]he presence of a third player who is penalized by a rejection should lead altruistic responders to accept lower offers“. Eigennützige Proposer könnten dies antizipieren und daraufhin geringere Angebote stellen. Tatsächlich zeigte sich Responder Y in Gegenwart von Z eher bereit, geringere Angebote zu akzeptieren.

Kagel und Wolfe (2001) spielten das Experiment von Güth und van Damme (1998) nach, und zwar mit der Änderung, dass der Responder pro Runde zufällig in der Rolle von Y oder Z ist, nachdem X seine Allokation gemacht hat, um Z nicht durch X zu marginalisieren und Y zu „kaufen“. Sie untersuchen damit die Änderung des Akzeptanzverhaltens, wenn wegen einer Zurückweisung ein inaktiver dritter Spieler unterschiedliche Gewinne zur Konsolidierung empfängt. Wenn also Y eine Offerte ablehnt, dann erhält Z einen positiven Betrag und es wird Einkommensungleichheit zwischen Y und Z hergestellt. Auch hier konnte keine Zurückweisung auf den geringen Anteil für Z zurückgeführt werden, so dass erneut das Desinteresse am dritten Spieler bekräftigt wird – ähnlich wie bei Güth und van Damme (1998), nicht aber mit Güth und van Damme aus dem Jahre 2007 sowie Knez und Camerer (1995). Sie vermuten, dass die Erklärung derartigen Verhaltens mittels negativer Reziprozität fruchtbar ist.

Bei Riedl und Výrašteková (2002) sollten zwei Responder gleichzeitig ihre Entscheidung über die Vorgaben des Proposers fällen. Die Zahl der akzeptierten möglichen Angebote war signifikant höher, wenn eine Ablehnung den anderen Responder positiv beeinflusst, als wenn der Responder negativ oder überhaupt nicht beeinflusst wird. Der Hälfte der Responder war ihre eigene Position im Vergleich mit der des Proposers und des anderen Responders wichtig, sie zeigten also soziale Präferenzen. Die andere Hälfte kümmerte sich wenig darum und erwies sich als recht egoistisch.

In der Studie von Okade und Riedl (2005) konnte der Proposer eine kleine oder große Koalition auswählen und danach Angebote vorgeben, welches Responder 1 (beziehungsweise Responder 2) ablehnen kann, falls er Koalitionspartner ist, ansonsten nicht. Sofern der eine Responder die Offerte annahm, konnte der andere Responder auch darüber entscheiden. Den Ergebnissen nach wird die Effizienz und Gleichheit von Koalitionsbildung anscheinend abhängig von gegenseitiger Fairness gemacht. Die überwiegende Mehrheit der Probanden war willens, andere Probanden unter eigenen Kosten von der Koalition auszuschließen und wählte damit ineffiziente Teilkoalitionen. Dieses Ergebnis ungleicher Ressourcenverteilung und sozialer Exklusion kam zustande, wenn die Responder sich reziprok und die Proposer sich wie rationale Nutzenmaximierer verhielten. Die Proposer boten in 3-Personen Koalitionen auf unfaire Weise pro Responder je 25 % an, was mit Reziprozität beantwortet wurde. Die Studie erhärtet die Ignoranz gegenüber dem dritten Spieler.

Bei Halko und Seppälä (2006) befanden sich die Spieler einmal in allen Positionen (1x Proposer, 2x Responder), die zwei Responder mussten einen simultanen Zug tätigen (ähnlich Fischbacher et al. 2009). Dabei wurden wegen anderer Aufteilungsregeln nun vermehrt kleinere Angebote akzeptiert. Auch hier sorgte die Situation der Responder für geringere Offerten seitens des Proposers, ebenso für häufigere Gleichaufteilungen und dennoch recht hohe Ablehnungsraten. Die Responder akzeptierten geringere Angebote, wenn die Wahl des anderen Responders die eigene Auszahlung nicht beeinflusste. Sie waren damit weniger gewillt, den Proposer zu bestrafen, falls sie durch die Annahme eine positive Auszahlung einwandfrei sicherstellen konnten.

Die wenigen bisherigen Studien zu Ultimatumspielen mit multiplen Respondern liefern durchaus gemischte Ergebnisse. Teilweise lassen sich die egoistischen Verhaltensweisen des homo oeconomicus bestätigen. Aber genauso häufig treten eher uneigennützige Akte auf, die entsprechende soziale Präferenzen offenbaren.

4 Das Experiment in Bern

Die studentischen Seminarteilnehmer des empirischen Forschungspraktikums entwickelten im Herbstsemester 2012 experimentelle Designs zur Überprüfung des Verhaltens der Probanden auf soziale beziehungsweise egoistische Präferenzen. Im Frühlingssemester 2013 wurden die Experimente durchgeführt. Im Folgenden sind zwei der Treatments, die an die bisherigen Studien anknüpften, dargestellt.

4.1 Forschungsdesign

In einigen Untersuchungen (Fehr und Fischbacher 2003, 2004) müssen Probanden eine Reaktion auf hypothetische Entscheidungen der Mitspieler abgeben. Fehr und Fischbacher (2004) legen nahe, dass darin eine Fehlerquelle liegen könnte. Das ist nach Ansicht von Elster (2005) höchstwahrscheinlich auch der Fall, und zwar weil die so hervorgerufenen Reaktionen (negative Emotionen) zu echten im Vergleich zu potenziellen Verletzungen der Fairnessnorm wesentlich stärker ausfallen können. Fehr unternimmt diese Experimente dennoch und zwar mit folgender Rechtfertigung: Da Diktatoren praktisch nie ganz konkrete (also im hiesigen Fall: zu niedrige) Transferniveaus auswählen, wären letztlich keine Daten für diese Level vorhanden, sofern nur Reaktionen auf echte Entscheidungen erlaubt wären (zitiert nach Elster 2005, S. 209). „This problem could be overcome by having subjects respond to computer-generated transfers at any level, as long as they thought they were dealing with a real person. Given the anonymity of experiments, this should be easy to achieve.” (Elster 2005, S. 209).

Die hier angewandten Ultimatumspiele beinhalten genau diese Änderung gegenüber anderen Studien, weil sie nämlich auf der fingierten unfairen Vorgabe (8 zu 2 beziehungsweise 9 zu 1) eines fiktiven, vorgetäuschten Proposers basieren, und zwar nach Vorbild des sequenziellen Diktatorspiels bei Diekmann (2004). Die Einhaltung der Fairnessnorm eines echten Spielers in der Rolle des Proposers ist damit irrelevant. Maßgeblich für die Untersuchung sind die Reaktionen der Responder in den verschiedenen experimentellen Gegebenheiten. Es wird unterstellt, dass die Responder die Fairnessnorm anerkennen und erwarten, dass der vermeintlich echte Proposer ihr ebenso Folge leistet. Das zuvor dargestellte einfache Ultimatumspiel mit nur einem Responder und denselben fingierten Vorgaben dient hierbei als Kontrollgruppe und lässt sich auch mit den Ergebnissen der Ultimatumspiele aus anderen Studien vergleichen.

4.1.1 Konkurrenzsituation

Das Design sieht vor, dass zwei Responder in besonderer Weise sequentiell, also nacheinander über ein- und dasselbe Angebot entscheiden. Zunächst erhält Responder 1 die Möglichkeit, den aufgeteilten Betrag anzunehmen oder abzulehnen. Sollte er sich für ersteres entscheiden, so erhalten er und der Proposer den jeweiligen Betrag ausgezahlt und das Spiel ist damit beendet. Responder 1 erhält keine Auszahlung, falls er das Angebot ausschlägt, und es wird automatisch an Responder 2 weitergeleitet. Dieser kann nun das bereits einmal abgelehnte Angebot ebenfalls annehmen oder ablehnen. Falls er das Angebot annimmt, erhalten er und der Proposer den aufgeteilten Betrag. Wenn Responder 2 ablehnt, erhält keiner der Spieler eine Auszahlung und das Spiel ist vorbei. Man kann sich diesen Teil des Treatments als einfaches Ultitumspiel vorstellen, der erste Responder hat (zumindest theoretisch) keinen Einfluss mehr auf die Entscheidung.[15] Bei dieser Art des Spiels herrscht also Konkurrenz unter den Respondern dergestalt, dass nur einer der beiden seinen Anteil bekommen kann, wie dies aus Abbildung 3 hervorgeht. Die Probanden spielen jeweils einmal die Positionen von Responder 1 und Responder 2, Lernprozesse sind damit irrelevant. Die Vorhersagen sind ähnlich denen des Mini-Ultimatumspiels, so dass bereits der erste Responder das unfaire Angebot annimmt, um sich selbst nicht schlechter zu stellen. Demnach dürfte der zweite Responder gar nicht zum Zuge kommen. Dennoch werden die Probanden in diese Position manövriert und daher sollte auch der zweite Responder das unfaire Angebot bestenfalls akzeptieren. Das Design wird hier als „Konkurrenzsituation“ bezeichnet, der Begriff ist aber unscharf und bisweilen irreführend, da die Responder mit dem Proposer im ökonomischen Sinne nicht echt kommunizieren und über den „Preis“ verhandeln können – es gibt keinen Markt mit zeitgleichem Bietkampf um die zulässige Unfairness. Treffender wären vermutlich emotionale Schlagworte wie „Missgunst“ oder „Neid“, denn es wird mit dem Arrangement darauf abgezielt, dass einem Responder 1 ein aggressives Angebot im Angesicht des Responders 2 als weniger unfair erscheint, als ohne letzteren.[16]

---

[15] Aus diesem Grund wird in der statistischen Analyse das Ablehnungsverhalten des zweiten Responders separat behandelt.

[16] Die in Bern dazu formulierten, hier nicht weiter verfolgten Hypothesen lauteten:

H1a: Die „Konkurrenzsituation“ veranlasst den ersten Responder zur Annahme des Angebots, um den Verlust der Zuteilung zu vermeiden.

H1b: In der Konkurrenzsituation sollte der zweite Responder das Angebot annehmen.

H1c: Die Ablehnungsraten sind unter Konkurrenz geringer als im normalen Ultimatumspiel.

Der zweite Responder kann hier als ein (zunächst) machtloser Dummyspieler gedacht werden (Z), dem nicht zwingend Anteile zukommen müssen und der in Konkurrenz zu dem ersten Responder steht (Y). Es kommt damit der Versuchanordnung von Güth und van Damme (1998, 2007) zwar nahe, ist aber wegen der tatsächlich aktiven Entscheidungsmöglichkeit nicht identisch.

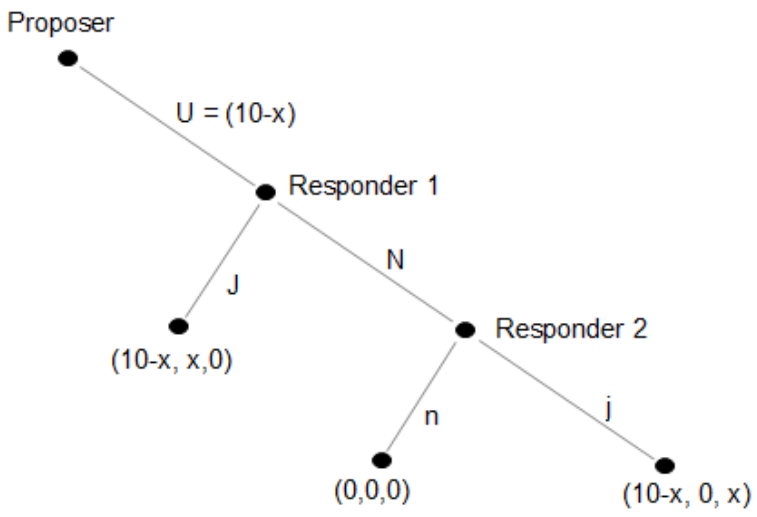

Abb. 3: Ultimatumspiel mit multiplen Respondern in sequentieller Anordnung (extensive Form). Die Höhe des Angebots des Proposers wird zufällig variiert, der angebotene Anteil x ist entweder 1 oder 2 (von 10 Franken), gleichermaßen in der Kontrollgruppe und im simultanen Design.

4.1.2 Verantwortungsdiffusion

Das zweite Design ist eine modifizierte Form des Ultimatumspiels, bei dem die Responder simultan und unabhängig voneinander über das Angebot des Proposers entscheiden. Die Variation besteht darüber hinaus in der Zahl der Responder – ein Teildesign besitzt zwei und das andere vier. Wenn einer der Responder die Auszahlung ablehnt, so erhalten sowohl dieser als auch der fiktive Proposer nichts, alle anderen annehmenden Responder erhalten jedoch ihren Anteil. Eine strafende Antwort an den Proposer, aber auch der Eigennutz kann jeweils unilateral verfolgt werden und ist in keiner Weise an Ablehnungen oder Zusagen anderer Spieler gekoppelt. Dabei stellt sich jedem einzelnen Responder die Frage, warum gerade er die kostspielige Sanktion des Proposers übernehmen soll und nicht jemand anderes. Es wird auf den Effekt einer Verantwortungsdiffusion beziehungsweise das sanctioning dilemma abgezielt. In Abb. 4 wird der Versuchsaufbau schematisch dargestellt. Anhand der Auszahlungskombinationen wird ersichtlich, dass der Anteil des Proposers von der Entscheidung des jeweils anderen Responders abhängt. Dieses Design kommt damit dem von Fischbacher et al. (2009) nahe. Die Unterschiede bestehen in der fingierten Vorgabe des Proposers und im Auslassen der Randomisierung der Auszahlung an annehmende Responder.

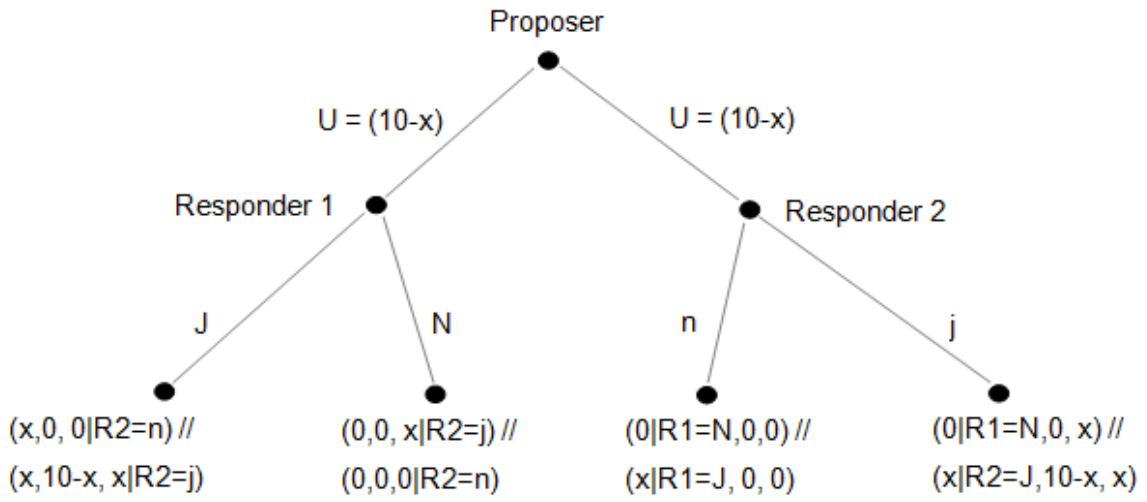

Abb. 4: Ultimatumspiel mit multiplen Respondern in simultaner Anordnung (extensive Form).
Es wird hier bereits ein Problem des Designs deutlich, denn falls alle zwei bzw. vier Responder das Angebot akzeptieren, summiert sich die Auszahlungskombination auf 12 bzw. 14 Franken, was eigentlich nicht möglich sein dürfte. In diesem Falle könnten clevere Responder die „Bank betrügen“.

Die beiden hiesigen Designs mit multiplen Respondern knüpfen bei den dargelegten, verwandten Studien an und versuchen eine Synthese verschiedener Elemente. Dennoch wird auf einige bisher angewandte Punkte verzichtet: So werden weder dem Proposer noch den Respondern „outside-options“ zur Verfügung gestellt wie bei Knez und Camerer (1995). Vor allem hinsichtlich des unfairen Angebots des künstlichen Proposers werden sie in einen Zwang versetzt. Häufig werden die Angebote in Geldeinheiten angegeben, über mehrere Runden verhandelt und anschließend in „echtes“ Geld umgetauscht. Hier spielen die Probanden jeweils einmal in der Rolle des Responders und des Proposers um die Aufteilung von jeweils 10 Franken anstatt um später zu konvertierende Währungen („Taler“).

Wie oben erwähnt, kann die Befolgung einer Norm teilweise strategisch motiviert sein. Das experimentelle Design versuht genau das so weit wie möglich zu verhindern. Es hemmt in seiner Auslegung bereits Effekte dieser intervenierenden Variablen, unter anderem aus mehreren Spielrunden gewonnene Lerneffekte oder soziale Erwünschtheit (Pointner 2012, S. 75). Sollte altruistic punishment einzig und allein durch einen „warm glow“ motiviert sein, kann sein Auftreten weder gemessen noch unterbunden werden.

4.2 Charakteristika des Designs

Die Versuchsanordnungen weisen inklusive Kontrollgruppe einige Besonderheiten auf. Zum einen der Rollentausch von der Responderposition hin zur Situation als Proposer mit schließlich echten Vorgaben natürlicher Personen an fiktive Responder. Zum anderen das Ausbleiben von Wiederholungen dieser „Spielrunde“. Daneben sind andere Merkmale wichtig, auf die nachstehend kurz eingegangen wird.

4.2.1 Das fingierte niedrige Angebot

Ein annähernd faires Angebot sollte bei zwei Respondern pro Spieler etwa 33.3 % betragen (einschließlich Proposer). Das hätten im Falle von 10°Franken also etwa 3.33°Franken sein müssen und für die Aufteilung von fünf Euro in der Reproduktion dementsprechend etwa 1.66°Euro pro Person. Die faire Aufteilung sähe im Falle von vier Respondern für jeden Spieler einen Euro vor. Eine Verletzung der gebotenen Fairness verspräche dem Proposer 4°Euro (beziehungsweise 3°Euro) und den Respondern jeweils 0.50°Euro (beziehungsweise 0.50°Euro). Da die meisten Spieler wie bereits dargelegt (Elster 2005) diese normalerweise eher scheuen, ist eine computergenierte ungleiche Verteilung des Proposers notwendig, um den eindeutigen Verstoß gegen die Norm der Fairness zu gewährleisten und die Reaktion zu provozieren.
Die Anordnung der Responder liefert zugleich ein theoretisch und empirisch fundiertes Argument dafür, dass die Aufteilung plausiblerweise so niedrig ausfallen kann. Denn je mehr Responder zugegen sind, desto geringer ist das Zurückweisungsrisiko für den Proposer. Fraglich ist eben, ob die Responder sich selbst in ihrer anschließenden Rolle als Proposer danach richten oder nicht.

4.2.2 Spielertypen

Man kann generell davon ausgehen, dass es in jeder Population von Probanden einen – mehr oder weniger hohen – Anteil an egoistischen Spielern und damit auch eigennützigen Respondern gibt. Eine Erhöhung der Spielerzahl erhöht zugleich auch die Zahl der Spielertypen. Je mehr Spieler also am Experiment teilnehmen, desto höher fällt die Wahrscheinlichkeit aus, dass mindestens einer der multiplen Responder egoistisch ist. Sie nehmen dann das unfaire Angebot an und teilen selbst unfair auf, vielleicht weil sie die Mitspieler für genauso eigensinnig wie sich selbst halten (Knez und Camerer 1995, S. 71).
Gleichzeitig werden sich darunter auch eher diejenigen Responder befinden, die Nutzen nicht nur aus Einkommen sondern auch aus Reziprozität ziehen, und zwar „[...] from reducing the earnings of those who fail to take actions that uphold the social norm“ (Denant-Boemont et al. 2007, S. 166).

Derartige stark reziproke Probanden (Fehr und Fischbacher 2003, S. 788) würden wohl das unfaire Angebot ablehnen und in einer direkten Austauschsituation den niedrigen zuvor angebotenen Betrag erwidern (Pointner 2012). Die Situation ist im vorliegenden Experiment aber nicht direkt, da eine zufällige Neuzusammensetzung lediglich suggeriert wird, was die multiplen Dyaden zwischen Proposern und Respondern also nicht identisch werden lässt. Akteure, die ihren sozialen Präferenzen folgen, würden daher trotz eigens zuvor erfahrener unfairer Behandlung wiederum faire Angebote tätigen. Denn es ist nicht gewährleistet, dass der von ihnen bestrafte Proposer nun auch als Responder mit einem ähnlich aggressiven Angebot konfrontiert werden kann.

Die Zahl der egoistischen Trittbrettfahrer und der altruistisch Bestrafenden korreliert positiv miteinander, was wie folgt erklärt wird: „When punishers are common defectors do badly, but when punishers are rare defectors do well“ (Boyd et al. 2003, S. 3532). Wie aus den zahlreichen Studien hervorgeht, ist etwa die Hälfte der Subjekte typischer Gewinnmaximierer, der Rest offenbart Präferenzen für Fairness (Güth und Kocher 2014, S. 7f.).

Eine weitere handlungsleitende Norm ist die der „conditional cooperation“ (Fehr und Fischbacher 2003, S. 788). Demnach tragen Akteure genau dann zu einem Kollektivgut bei, wenn andere oder die Mehrheit es auch tun. Im Hinblick auf die Defektion der anderen stellt dies eine legitime Rechtfertigung der Defektion durch Ego dar. Vermutlich üben einige Responder auch zuerst eine Sanktion aus, weil sie damit rechnen, andernfalls mit einer gewissen Wahrscheinlichkeit in der nächsten Runde bestraft zu werden (Denant-Boemont et al. 2007, S. 166). Da diese in der nächsten Runde ausbleibt, wenn sie selbst ein Angebot offerieren, könnten sie ihre Strategie von altruistisch hin zu eigennützig wechseln. Eventuell manifestiert sich also noch eine Mischform solcher Akteure, die unfaire Angebote ablehnen und selbige danach anbieten.

4.2.3 Die Spielstruktur

Fehr et al. (2002, S. 54) wiesen allgemein darauf hin, dass die angenommenen Angebote bei mehreren Spielern durchaus niedriger ausfallen können und auch äußerst unfaire Angebote (Anteile von 10 %) hingenommen werden, obwohl sie im klassischen Ultimatumspiel womöglich abgelehnt worden wären.

Unsichere Umstände in Anbetracht mehrerer Akteure könnten zu rationalerem Verhalten führen. Es ist möglich, dass egoistischere Spielertypen dann den Eigennutz-Modus ausspielen anstatt strategisch motiviert prosoziales Handeln zu signalisieren. Fraglich ist hingegen, ob Spieler mit eher kooperativen sozialen Motiven verstärkt abwägen in ihren Entscheidungen. Allerdings gelten diese Zusammenhänge wohl vorerst eingeschränkt, denn unter Unsicherheit wurde selbst die effiziente Strategie des tit-for-tat kompromittiert (Weber et al. 2004,S. 291f.).

Die Struktur des Spiels kann also dazu führen, dass egoistische Akteure ihr altruistisches Pendant dominieren (Fehr und Gächter 2002, S. 137; Fehr und Fischbacher 2002, C4), et vice versa: “Depending on the environment, a minority of altruists can force a majority of selfish individuals to cooperate or, conversely, a few egoists can induce a large number of altruists to defect” (Fehr und Fischbacher 2003, S. 785). Folglich akzeptieren – eigentlich vielleicht prosozial eingestellte – Responder häufiger die angebotenen Aufteilungen, wenn sie mit vielen anderen Respondern konkurrieren, anstatt zu defektieren. Obschon das Signal also egoistisch sein kann, ist ihre Gesinnung eigentlich sozial und reziprok: „Das Ergebnis einer Handlungswahl ist manchmal nicht fair […], hingegen können die zugrunde liegenden Intentionen des Mitspielers aber fair sein“ (Pointner 2012, S. 50).

4.3 Durchführung

Die Stichprobe wurde aus der Grundgesamtheit der gesamten Studentenschaft der Universität Bern gezogen. Die potenziellen Teilnehmer erhielten ein Einladung in den PC-Pool des Instituts für Soziologie der Universität Bern (Gebäude des UniToblers). Dort fand im April 2013 die Erhebungsphase statt, die auf insgesamt 12 Sessions á höchstens 30 Minuten verteilt war. Die Gruppengröße variierte pro Sitzung zwischen 3-15 Probanden. Zwischen den Computern war an jedem Sitzplatz ein Sichtschutz aufgebaut. Anschließend wies der Experimentalleiter auf die Vorgehensweise, den Ablauf und die Regeln hin und beantwortete etwaige Fragen.
Das Experiment wurde mit der Software z-Tree (Fischbacher 2007) programmiert und durchgeführt. Die Entscheidungssituationen und Anweisungen wurden am Bildschirm schrittweise erläutert[17]. Nach diesen Spielen wurden die Probanden aufgefordert, z-Leaf zu beenden und im Browser einen Online-Fragebogen auszufüllen. Mit diesem wurden Merkmale der Sozialstatistik, Netzwerkkontakte und Vertrauen (in fremde Menschen und Institutionen) erhoben. Die Computer waren mit nummerierten Zetteln versehen, die die Teilnehmer an der Auszahlungsstelle abzugeben hatten. Die erspielten Endbeträge wurden durch den Leiter in Form einer Excel-Tabelle per E-Mail zur Auszahlungsstelle im Nachbarraum geschickt, wo das Geld abgezählt und in Kuverts bereitgestellt wurde. Eine für diesen Zweck angefertigte Tür hatte eine kleine Öffnung und konnte nur von innen geöffnet werden. Wenn die Probanden ihren individuellen Zettel durch diesen Briefkastenschlitz hindurch steckten, wurde der entsprechende Umschlag mit dem erspielten Geldbetrag retour ausgegeben. Somit war auch die Anonymität der Auszahlung sichergestellt.

[17] Genaue Formulierungen finden sich im Anhang beziehungsweise auf der externen, beigefügten CD.

4.4 Ergebnisse

4.4.1 Univariat

Am normalen Ultimatumspiel nahmen 36 Probanden teil, 42 konnten die sequentielle Konkurrenzsituation durchspielen, während sich 39 im simultanen Design befanden. Die insgesamt 117 Teilnehmer kamen fast ausnahmslos von der Universität Bern und verteilten sich auf verschiedene Fakultäten: Knapp ein Fünftel kam von der Wirtschafts- und Sozialwissenschaftlichen Fakultät, ein Viertel von der Philosophisch-historischen Fakultät, 28.4 % von der Vetsuisse und der Rest verteilte sich auf die übrigen. Es gab 28 Studenten der Wirtschaftswissenschaften (Haupt- oder Nebenfach), von denen 55.56% männlich waren. Nach eigenen Angaben haben knapp 77 % der Personen noch nie an einem derartigen Experiment teilgenommen. Vom Rest waren lediglich 3 Personen durch häufigere Experimente mit der Thematik vertraut. Es beteiligten sich 73 Frauen (64.04 %) und 41 Männer (35.96 %)[18], zusammengefasst im Schnitt 23.7 Jahre alt. Das monatliche Nettoeinkommen der Studenten lag bei durchschnittlich 617.48 Franken mit einer Streuung von 0 bis 3800 Franken. Niemand verweigerte die Angabe.

Aufgegliedert nach Designs entschieden sich etwa 55.6 % der Probanden im einfachen Ultimatumspiel dafür, die angebotenen ein beziehungsweise zwei von 10 Franken auszuschlagen. Folglich findet dieses Ergebnis Anschluss an die oben dargelegten Untersuchungen, die Zurückweisungen von etwa 50 % für solche Angebote dokumentieren. Die Ablehnungsrate beträgt in der Konkurrenzsituation 73.81 %, wobei sie in der als zweiter Responder mit 78.57 % nicht signifikant höher liegt ($t = -1.00$). Im simultanen Design fiel sie mit 74.36 % praktisch gleich hoch aus. Trotz der hohen Ablehnungsraten ist noch ausreichend Verhaltensvarianz vorhanden, die Probanden agieren nicht ausschließlich fair, sondern zum Teil auch egoistisch.

Die eigens gestellten Angebote im sequentiellen Design fielen mit im Durchschnitt 3.83 Franken tendenziell fair aus. Es gab keine hyperfairen Angebote von mehr als fünf Franken und auch dieser Befund schließt an bisherige Studienergebnisse an (Camerer 2003, S. 49). Die Werte sind mit denen des simultanen Designs nahezu identisch, so dass keine signifikanten Unterschiede zwischen beiden Treatments bestehen ($t = 0.054$). Sie liegen außerdem nur geringfügig unter denen aus der Kontrollgruppe mit durchschnittlich 4.01 Franken (Median=5).

Der neben der Aufwandsentschädigung von 20 Franken erspielte Verdienst gestaltete sich wie folgt: Im Mittel erspielten die 81 Personen der Treatments 6.28 Franken (Median bei 6). In diesem Punkt gab es zwischen den beiden Designs mit multiplen Respondern offenbar keine signifikanten

[18] In der Summe sind das nur 114 statt 117 Personen. Offenbar haben sich drei Probanden dieser Kategorisierung verweigert. Weiterhin sind im simultanen Design 76% weiblich, in den anderen 57%.

Mittelwertunterschiede (Differenz=0.045, t = 0.1147).
Über den Fragebogen im Anschluss an das Experiment wurden weitere Variablen erhoben, unter anderem zur Soziodemographie. Die angegebenen minimal akzeptablen Angebote (folglich MAO) lagen im Mittel bei 3.20 Franken, mit einem Median bei 4 Franken. Es manifestiert sich also eine Tendenz zur Forderung einer fairen Aufteilung[19]. Das normale Ultimatumspiel wurde mit durchschnittlich 6.61, das simultane mit 6.54 und das sequentielle mit 5.98 Punkten als überwiegend anonym eingeschätzt. Die Probanden in der Konkurrenzsituation fühlten sich offenbar signifikant weniger anonym als die in der Kontrollgruppe (t = -2.64) beziehungsweise als die im Freiwilligendilemma (t = 2.18).

4.4.2 Bivariat

Der erwartete Effekt der Anwesenheit multipler Responder, geringere Ablehnungsraten als in der Kontrollgruppe zu produzieren, stellte sich nicht ein. Mithin sind die Unterschiede beider Designs mit multiplen Respondern im Vergleich zum normalen Ultimatumspiel auf 5%-Niveau knapp nicht signifikant (t= -1.70 beziehungsweise t= -1.72).
Es manifestiert sich eine Auswirkung der Anwesenheit mehrerer Responder auf die Neigung zur Ablehnung des Angebots (also 81 von insgesamt 117 Personen in erster Responderposition). Allerdings dergestalt, dass die Tendenz zur Ablehnung erhöht wird (r=0.18**), obschon sie aufgegliedert nach Designs nicht bestehen bleibt (simultan: r=0.09 und für den ersten beider sequentiellen Responder: r=0.09, gerundete Werte). Der Effekt ist zwar nicht sonderlich stark, deutet dennoch zweimal in dieselbe Richtung.
Es ließen sich keine statistisch signifikanten Unterschiede ermitteln abhängig davon, ob die fingierte Vorgabe 9:1 oder 8:2 war. Das kostengünstigere altruistic punishment wird nicht häufiger ausgeübt. Offenbar hat die geringfügige Varianz in der Aufteilung des Betrags keinen Einfluss auf die Entscheidungen der Responder – es erscheint so oder so als sehr unfair.
Ein wesentlicher Unterschied in den Ablehnungsraten zwischen den Frauen (68.5 %) und Männern (65.8 %) ließ sich nicht feststellen, so wie er in obigen Studien teilweise beobachtbar war. Ein Geschlechtereffekt trat mithin nicht auf.
Die Entscheidungen der Probanden, die beide Responderpositionen im sequentiellen Design durchspielten, waren äußerst konsistent. Insgesamt entschieden sich nur vier Personen, ihre

[19] Da die Variable zeitlich nach den Entscheidungen mit monetären Konsequenzen fiel, ergibt sich ein Methodeneffekt, der relevante Variablen in der der ursprünglichen multivariaten Analyse bei weitem überwiegt und daher folglich statistisch und inhaltlich vernachlässigt wird.

Meinung zu ändern. Drei nahmen das Angebot das erste Mal an und das zweite Mal nicht, nur eine Person wechselte in die Gegenrichtung. Die Probanden sind also recht konsequent in ihrer Orientierung an der Fairnessnorm gewesen.

Über alle 117 Probanden zusammengefasst ergibt sich eine mittlere Vorschlagshöhe von 3.89 Franken. Darunter geben die 73 Frauen etwa 0.52 Franken mehr ab als Männer (t=2.2). Jene Probanden, die das erste unfaire Angebot akzeptierten (also 37 von 117), geben selbst mit durchschnittlich 3.14 Franken signifikant 1.11 Franken weniger ab (t= 5.08) als diejenigen, die das Angebot ablehnten (80 von 117 Personen). Nach den Designs unterteilt manifestieren sich diese Unterschiede für das normale Ultimatumspiel (1.60 Franken weniger, t =3.68) und für das Freiwilligendilemma zumindest auf 10%-Niveau (47 Rappen weniger, t=1.26). Gleiches gilt unter Konkurrenz für den ersten Responder (1.38 Franken weniger, t = 4.22) und den zweiten Responder (1.77 Franken weniger, t = 5.65). Nicht einmal 10 % der Probanden geben genauso viel oder weniger ab, als sie vom fingierten Proposer angeboten bekamen. Anscheinend halten die meisten Personen weiterhin an der Fairnessnorm fest, obgleich sie selbst mit einem aggressiven Angebot konfrontiert wurden. Sie lassen nicht grundlos (mindestens) einen Mitspieler ein unfaires Angebot erleiden, der nicht für die erstmalige unfaire Aufteilung verantwortlich ist – sie verhalten sich in der Regel weder direkt, noch indirekt reziprok.

Weiterhin hat sich gezeigt, dass die Ablehnungsrate der 28 Wirtschaftswissenschaftler über beide Treatments mit jeweils knapp 90 % geringfügig höher ist als bei den übrigen Studierenden (etwa 70 %). Den Befund muss man nicht überbewerten, da der Zusammenhang schwach und nicht signifikant ist und die Fallzahl ohnehin sehr klein. Selbige Teilpopulation erspielt insgesamt 0.98 Franken weniger (t=2.17), obschon sich deren um durchschnittlich 12 Rappen niedrigere Vorschläge nicht von denen der anderen Studenten unterscheiden (t=0.44). Demnach haben sie nicht weniger Verdienst erhalten, weil sie mit egoistischen Vorgaben zu riskant oder aggressiv waren. Ein Effekt des Faktors Studienfach lässt sich bivariat nicht belegen. Auch im Hinblick auf eine Interaktion mit dem Geschlecht gibt es keine Evidenz, da nur 12 Studentinnen Wirtschaft im Haupt-oder Nebenfach belegen.

Das angegebene Anonymitätsempfinden der Probanden lag über alle Designs mit durchschnittlich 6.36 Punkten sehr hoch. Es hatte zwar keinen Einfluss auf das Ablehnungsverhalten, aber in der Kontrollgruppe fühlten sich ablehnende Personen um 0.425 Punkte anonymer (t=-2.05). In der Konkurrenzsituation war es um 0.56 Punkte geringer als im Freiwilligendilemma (t = 2.18) und um 0.64 Punkte geringer als in der Kontrollgruppe (t=-2.64), zwischen den letzteren bestanden keine signifikanten Unterschiede. Anonymität förderte nur in der Kontrollgruppe geringfügig die eigene Vorschlagshöhe (r=0.32*) und kommt hier offenbar nicht als Fairness beeinflussende Störgröße in Betracht.

4.4.3 Multivariat

Die hier präsentierten Ergebnisse sind in Anbetracht des Umfangs nicht besonders detailreich. Die Variable „multiple Responder“ ist zusammengefasst, um die multivariate Analyse einfacher und übersichtlicher zu halten, zumal beide Teildesigns keine statistisch signifikanten Unterschiede in den Ablehnungsraten aufweisen. Die kontraintuitiven Tendenzen beziehungsweise das Ausbleiben der Treatment-Effekte der Ergebnisse sind bereits in den vorangegangenen Abschnitten ersichtlich[20].

Eine Regression der Ablehnungen der fingierten Angebote auf das Treatment multipler Responder als Linear Probability-Modell (folglich LP-Modell) spiegelt die vorigen Ergebnisse wider: Die Konstante liegt bei 0.56***, also der Ablehnungsrate in der Kontrollgruppe, und der Koeffizient des Treatment-Effekts liegt bei 0.18***, er erhöht also die geschätzte Wahrscheinlichkeit der Ablehnung, ganz so wie die Ergebnisse der Versuchsgruppen.

In einer logistischen Regression erhöht die Anwesenheit von multiplen Respondern die vorhergesagte Chance auf Ablehnung des Angebots um den Faktor 2.29, eine signifikante Verdopplung. Fast identische Werte erzielt alleinig der Effekt der eigenen Abgaben. Jeder zusätzlich abgegebene Euro erhöht die vorhergesagte Chance signifikant um den Faktor 2.24. Wenn diese beiden Haupteffekte verbunden werden (Modell 3), dann erhöhen sich ihre absoluten Werte, der Einfluss des Treatments ist deutlich stärker und steigt auf 3.61 an. Die Effekte beider sind gemäß Likelihood-Ratio-Test (folglich LR-Test) in der Grundgesamtheit sehr wahrscheinlich nicht null und verbessern die Schätzungen im Vergleich zu Modellen ohne die jeweils andere Komponente signifikant.

Sofern nun noch die Variable des wirtschaftswissenschaftlichen Studiums hinzugefügt wird, steigen die Effekte der zentralen Variablen nochmals an, diesmal etwas moderater. Der Effekt des Wirtschaftsstudiums fällt mit 4.03 allerdings stärker aus als der des Treatments mit 3.87 oder der des eigenen Vorschlags mit 2.83, alle drei sind für sich signifikant. Dieses Modell ist noch etwas besser als seine Vorgänger, der Einfluss der aufgenommenen Variable ist wahrscheinlich nicht null und sollte daher bei der Schätzung nicht vernachlässigt werden.

[20] Der Zusammenhang des Regressoren „eigener Vorschlag“ mit der abhängigen Variablen ist nicht perfekt linear (steigend), aber zumindest im Schwerpunkt der Daten und auch im Wertebereich von 0.2 bis 0.8 (Kohler & Kreuter 2008, S. 284f.). Ebenso ist dies der Fall im Definitionsbereich bis zur fairen 50-50-Grenze, es existieren keine Beobachtungen hyperfairen Bereich. Gleiches gilt im Übrigen für die Analyse der Daten der Reproduktion weiter unten.

Abhängige Variable: Ablehnungsverhalten (erster Responder), Bern 2013

Unabhängige Variablen	**Korrelation mit AV**	**Modell 1**	**Modell 2**	**Modell 3**	**Modell 4**	**Modell 5**
Haupteffekte						
Multiple Responder	0.18**	2.29**		3.61**	3.87**	4.29**
eigene Abgaben	0.43***		2.24***	2.54***	2.83***	2.97***
Kontrollvariablen						
Wirtschaftsstudium	0.12				4.03**	4.12**
fingiertes Angebot						0.80
Weibliches Geschlecht	0.03					0.67
Alter	0.01					0.90
Einkommen	0.01					1.00
Anonymitätsgefühl	-0.06					1.20
Teilnahme	0.07					1.60
Konstante		1.25	0.10***	0.03***	0.01***	0.03
Pseudo- R^2		0.03	0.15	0.19	0.23	0.24
BIC		-405	-432	-424	-425	-383
(Adjustiertes) Count-R^2		0.68 (0.00)	0.78 (0.30)	0.78 (0.30)	0.74 (0.19)	0.78 (0.32)

Abb. 5: Ablehnungsverhalten der Responder in Bern.

Logistische Regression auf die Entscheidung, das unfaire Angebot abzulehnen (abhängige Variable, Wert 0= akzeptiert, 1= nicht akzeptiert), exklusive der Entscheidung des zweiten Responders im sequentiellen Treatment. Angabe der Koeffizienten in Odds-Ratios. Signifikanzniveaus: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Das Pseudo-R^2 ist das von McFadden. Für die Modelle ohne Geschlechtervariable N=117, mit dieser N=114. Kodierung multiple Responder: Wert 0= Kontrollgruppe, 1= mit multiplen Respondern (n=81); Variable fingiertes Angebot: Wert 0= 9 zu 1°Franken, 1= 8 zu 2°Franken.

Das Modell 5 mit sämtlichen Kontrollvariablen ändert im Wesentlichen nichts an der vorhergesagten Chance der Ablehnung. Es erhöht jeweils etwas den Wert für das Treatment auf 4.29**, den der eigenen Abgaben auf 2.97*** und den des Wirtschaftsstudiums auf 4.12**. Die Odds-Ratios dieser Variablen liegen allesamt sehr dicht bei 1, sind für sich selbst relativ klein und auch wesentlich schwächer als die anderen drei wichtigen Haupteffekte. Sie deuten auf keinen Zusammenhang hin, ganz ähnlich ihrer bivariaten Korrelation mit der Ablehnungsrate. Da sie weder einzeln noch zusammen nichts sonderlich zur Varianzaufklärung beitragen, können sie (anhand LR-Tests) außen vor gelassen werden. Dennoch zeigen sie die relative Robustheit der zentralen Effekte aus den vorangegangenen Modellen.[21]

Die Modelle 1 bis 3 wurden zur Kontrolle in LP-Modellen gerechnet, die Koeffizienten sind für die eigene Vorschlagshöhe 0.17***, den Treatment-Effekt 0.22** und das Wirtschaftsstudium 0.18** - sie bestätigen damit die Effektrichtungen. Alle drei Variablen verbessern einzeln die Schätzung signifikant gemäß schrittweisem (incremental) F-Test. Die VIF-Werte liegen für alle unter 2, auch für das Vollmodelle mit sämtlichen Kontrollvariablen.

Das Ablehnungsverhalten als zweiter Responder der sequentiellen Entscheidung erweist sich multivariat genauso wie bivariat als deckungsgleich mit der Entscheidung als erster Responder. Die zuvor starke Korrelation beider (abhängigen) Variablen ist mit 0.75*** sehr hoch. Der Wert des Odds Ratio liegt bei 80*** (Standardfehler allerdings bei 97) und der des Koeffizienten im LP-Modell bei 0.69***. Unter Hinzunahme der eigenen Vorschlagshöhe sinkt der Odds-Ratio-Wert auf 64 (im LP-Modell auf 0.50). Der Effekt der eigenen Abgaben fällt noch stärker als in den bisherigen Modellen, der Odds-Ratio-Wert liegt bei 8.60*** und der Koeffizient in einem analogen LP-Modell bei 0.25*** (punktbiseriale Korrelation r=0.67***). Diese Variable ist essentiell und verbessert die Schätzung gemäß LR-Test signifikant, die typischen Kontrollvariablen hingegen nicht, sie sind demnach vernachlässigbar in ihrer Bedeutung.

Die Ergebnisse zu diesem Teildesign deuten in eine ähnliche Richtung wie die anderen logistischen Regressionsmodelle, sollten aber wegen der geringen Fallzahl nicht überbewertet werden. Zudem sind Aussagen zum Effekt der multiplen Responder nicht möglich (perfekte Multikollinearität) und wie bereits erwähnt inhaltlich als normales Ultimatumspiel zu interpretieren.

---

[21] Die Pearson-Residuen der bisherigen Regressionsmodelle sind gemäß Shapiro-Francia- und Shapiro-Wilk-Test nicht normalverteilt. Die meisten weisen eine Ballung um den Wert 0 auf, einige wenige Abweichungen verlaufen bis (-6) oder +3. Für alle Modelle existieren jeweils einige einflussreiche Fälle, für die die Zusammenhänge offenbar schlicht nicht zutreffen. Weil dies nicht stets dieselben Personen sind und die Stichprobe ohnehin nicht besonders groß ist, werden sie aus den Rechnungen nicht ausgeschlossen.

Entgegen den Erwartungen sind die Ablehnungsraten bei beiden Designs mit multiplen Respondern recht hoch. Multivariat ebenso wie zuvor bivariat zeigt sich: Die Anwesenheit multipler Responder veranlasst anscheinend weder im simultanen, noch im sequentiellen Design (erste Responderposition) zur Annahme des Angebots gemäß theoretischen Vorüberlegungen. Dieses Ergebnis reiht sich nicht bei den bisherigen Studien ähnlichen Aufbaus ein und liegt deutlich über dem Wert der Kontrollgruppe, obschon der gegenteilige Fall vermutet wurde. Die theoretischen Hypothesen konnten empirisch nicht gestützt werden. Weder die Gegenwart mehrerer Responder noch geringere Kosten für das altruistic punishment lassen letzteres häufiger vorkommen. Die Akteure verhalten sich gegenüber ihren Mitspielern dennoch recht fair.

4.5 Probleme des Experiments und Gründe für die Reproduktion

Diverse praktische Mängel sind bereits bei den Probeläufen im Vorfeld der Experimente aufgetreten. Diese wurden mit Experten (Dozenten und Studenten des Seminars) durchgeführt, welche durch die Gruppenreferate und Diskussionen im Seminar schon viel Vorwissen besaßen. Trotz größtmöglicher Einfachheit der Versuchsanordnung und größtmöglicher Antizipationsleistung etwaiger Probleme sind daher nicht alle Verständnisfragen und Hilfslosigkeiten in der jeweiligen Spielsituation bei den Experimentalteilnehmern ausgeräumt.

Bei der Analyse der durchgeführten Experimente fällt weiterhin die ungünstig niedrige Anzahl an Probanden auf. Denn gleich zweifach negativ wirkte sich die Tatsache der geringen Teilnehmerzahl aus. Zum einen insofern, als dass die Gruppengröße in einzelnen Sessions äußerst klein war (<<10) und damit vermutlich das Anonymitätsempfinden der wenigen Teilnehmer beeinflusste. Obschon die bivariate Auswertung zeigte, dass selbiges Gefühl bei einigen Probanden gering war, besteht jedoch kein signifikanter Zusammenhang mit dem Ablehnungsverhalten der Responder sowohl global als auch nach Designs. Zum anderen musste das Experiment mit vier Respondern zur Verantwortungsdiffusion ausgelassen werden, um wenigstens für jenes mit zwei Respondern genügend Probanden zu erhalten. Ohne dieses Aufstocken wären für jedes Teildesign zu wenige Versuchspersonen vorhanden gewesen.

Die Annahme zum abschließenden Fragebogen war, dass er wohl unterschiedlich schnell beantwortet würde. Damit sollte sichergestellt sein, dass sich keine Warteschlangen vor der Auszahlungsstelle bildeten, wobei die Probanden die Möglichkeit gehabt hätten, sich über das Experiment auszutauschen. Wenngleich die Aufteilungen dadurch nicht beeinträchtigt werden, so

ist eine Unterbindung dessen angebracht[22]. Leider war dies nicht stets gewährleistet. Überdies hatte der Fragebogen beinahe den Umfang einer eigenen standardisierten Befragung und ist damit deutlich zu lang für Laborexperimente.

Da die Ablehnungsrate weder in der Konkurrenzsituation noch im Freiwilligendilemma unter die Marke der Kontrollgruppe sinkt, stellt sich die Frage, ob die Teilnehmer die Darstellung der Situation nicht vollständig verstanden haben oder ob der Einfluss der multiplen Responder weniger stark ist als erwartet. Bezüglich letzterem ist die Interpretation naheliegend, dass das altruistic punishment in Anbetracht einer festen Teilnahmevergütung von 20 Franken mit nur einem beziehungsweise zwei von zehn Franken äußerst kostengünstig ausfiel. In dieser Hinsicht muss in der Reproduktion nachjustiert werden.

Ein Problem methodischer Art war die Zeitanzeige (Countdown) auf dem Bildschirm, wie viel Zeit es noch bis zur Entscheidung des Mitspielers bedürfe. Diese lief für alle Probanden gänzlich bis Null herunter, so als ob die jeweiligen Mitspieler sich genau mit Ende des Countdowns entschieden hätten und nicht variabel nach 18 oder 44 Sekunden. Erst danach ging die Tauschsituation weiter. Der eigentliche Zweck der Anzeige lag in dem Vortäuschen zunächst eines echten Proposers, der einer der anderen Probanden im Raum sei. Dies wurde schließlich nicht konsequent und raffiniert genug umgesetzt. Da währenddessen keinerlei Geräusche von den Teilnehmern des Experiments zu vernehmen waren, ist davon auszugehen, dass die Anzeige die Aufmerksamkeit noch erhöht hat und einige diese schlecht getarnte Falle durchschaut haben und ihr Entscheidungsverhalten änderten[23].

Ein etwas gewichtigerer Fehler trat in der Formulierung der Tauschlogik bei der Spielanleitung auf. Im Treatment sequentieller Responder wurde die Situation genau betrachtet so formuliert, als könnte der Proposer zweimal jeweils zehn Franken aufteilen, also insgesamt 20 Franken. Freilich wurde dies so nicht explizit mitgeteilt und erst recht nicht umgesetzt. Wenn man berücksichtigt, dass die Probanden in ihrer Rolle als zweiter Responder nicht wussten, ob der Proposer immer noch derselbe ist oder der andere Responder nun diese Position innehat, dann hätte der fingierte Proposer zweimal zehn Franken pro Responder aufteilen können, also viermal einen überproportionalen Anteil verdient.

---

[22] Es kann nur spekuliert werden, wie schnell sich die vermeintliche „Falle“ im Experiment herumspricht und die Teilnehmer der nächsten Session davon in Kenntnis setzt.

[23] Dies trotz der vorbeugenden Maßnahme, dass hin und wieder einige Seminarteilnehmer sich in den PC-Pool setzten, um durch reges Treiben an anderen Computern die herrschende Stille durch eine realistischere Geräuschkulisse auszutauschen.

Fehler in der Formulierung der Tauschlogik traten auch im simultanen Design auf. Zunächst wurde allen Respondern zwar versichert, dass für die altruistische Bestrafung des Proposers bereits ein einzelner ablehnender Responder ausreichen würde. Nach dem Rollentausch war die Anweisung als Proposer allerdings umgekehrt: Es würde bereits ein annehmender Responder ausreichen, damit der eigenen Proposeranteil sichergestellt sei. Dies ist ein inhaltlicher Fehler, der trotz Prüfung durch Dozenten und Studenten den Pretest passieren konnte. Vielleicht wurden daher manche (aufmerksame) Probanden in ihrer Entscheidung verunsichert.

Einen zusätzlichen Punkt bringt Diekmann (2008) an, indem davon auszugehen ist, dass viele publizierte signifikante Forschungsergebnisse fehlerhaft sind und das sogar bei korrekter statistischer Auswertung – was auch oft nicht der Fall sei. Alle bisherigen Effekte könnten reine Zufallsbefunde sein und besitzen demnach zunächst nur den Status einer Hypothese. Signifikante Effekte sollten daher durch weitere Studien reproduziert sein, bevor sie als robuste Nennwerte akzeptiert werden (ebd.). Im Hinblick auf den Beta-Fehler[24] können mögliche falsch-positive und falsch-negative Resultate durch Reproduktionen mit größerer Wahrscheinlichkeit verhindert werden. Um also überhaupt den Effekt der Verantwortungsdiffusion zu untersuchen (Erhöhung von zwei auf vier simultane Responder) und zweifelhafte Ergebnisse besser zu bewerten, empfiehlt sich eine Wiederholung des Experiments in Bern.

[24] Der ß-Fehler ist die Wahrscheinlichkeit ß, durch die unberechtigte Beibehaltung der Nullhypothese mittels Stichprobe die Alternativhypothese in der Grundgesamtheit zu verwefen, obschon sie eigentlich gilt. Diese ist unter anderem abhängig von dem Gegenpol alpha (Signifikanzniveau, gewöhnlich 0.05), der Fallzahl und der Effektstärke – also Faktoren, die hier eine Reproduktion nahelegen.

5 Hypothesen

Wie oben im Abschnitt zum symmetrischen Freiwilligendilemma angedeutet, ergeben sich durch die Modellierung (Diekmann 1983)[25] folgende Hypothesen zum Design der Verantwortungsdiffusion hinsichtlich des altruistic punishment, die in Bern teilweise identisch lauteten:

H1:
Die Verantwortungsdiffusion veranlasst die Responder individuell zur Annahme des Angebots, um die Kosten des altruistic punishments anderen zu überlassen.

H2:
Je größer die Anzahl simultan agierender Responder, desto geringer die individuelle Neigung zur Ausübung des altruistic punishment. Die Ablehnungsrate fällt bei 2 Respondern höher aus als bei 4 Respondern.

H3:
Die Ablehnungsrate liegt im simultanen Design tiefer als im normalen Ultimatumspiel. Je mehr Responder teilnehmen, desto größer fällt die Differenz zum Ultimatumspiel aus.

H4:
Je günstiger das altruistic punishment ist, desto häufiger wird es ausgeübt. Aufteilungen von 8 zu 2 werden häufiger angenommen als solche von 9 zu 1.

[25] Bartling & Fischbacher (2011) stellten die Hypothese auf, dass mit steigender Verantwortlichkeit eines Spielers auch seine Ausübung der Sanktion zunimmt. Bei dem hier verwendeten Design ist dies gerade gegenteilig, daher fügt sich die in Negation formulierte Hypothese in das Konzept ein. Vergleichbares findet sich auch bei Fischbacher et al. (2009, S. 540).

6 Reproduktion

Das Design zur „Konkurrenzsituation“ wies in Bern die meisten Fehler und Schwierigkeiten der Umsetzung auf. Die Kontrollgruppe zum normalen Ultimatumspiel konnte mangels Ressourcen nicht erneut erhoben werden. Die Reproduktion in Leipzig widmete sich daher nur dem Treatment zur Verantwortungsdiffusion. Vermutlich reicht bereits dieses eine Design aus, um die aufgetretenen Probleme abzustellen und Ergebnisse neu zu bewerten.

6.1 Modifikationen des Experiments gegenüber Bern

Die originären z-tree-Dateien aus Bern wurden vorab geringfügig geändert. Die Formulierungen der Entscheidungssituationen wurden berichtigt und zum besseren Verständnis gekürzt[26]. Daneben mussten zusätzlich die Beträge und Währungen angepasst werden. Um die Täuschung der Probanden zu optimieren, wurde jeweils vor den Entscheidungen eine kleine Quizfrage[27] eingefügt. Für die eine Teilgruppe der Responder geschah dies noch vor Beginn der ersten Offerte des fingierten Proposers, so dass die restlichen Responder abwarten mussten und durch die Eingabegeräusche überlistet wurden. Letztere bekamen dieselbe Frage, bevor sie ein echtes Angebot stellen konnten, so dass erstere getäuscht wurden. Den Probanden war versichert, dass die falsche oder korrekte Beantwortung keine Auswirkung auf ihre finanzielle Ausstattung besitzt. Die Anzeige „verbleibende Zeit“ (Countdown bis zur Entscheidung der Mitspieler) auf dem Entscheidungsbildschirm der Responder wurde bei der Reproduktion entfernt.
Die Tauschlogik ist sowohl in Bern als auch in Leipzig dieselbe, und zwar wird bei der Annahme des Angebotes seitens der Responder der „Kuchen“ größer. Es wird nicht jeder individuell angenommene Responderbetrag vom Kuchen abgezogen, so dass der Anteil des Proposers kleiner wird. Letzterer (beziehungsweise das unfaire Angebot) bleibt konstant und unabhängig davon wird jeder Anteil des annehmenden Responders hinzuaddiert. Ein Beispiel: Von insgesamt fünf Euro will der Proposer vier für sich behalten. Er bietet beiden Respondern je einen Euro an. Wenn beide annehmen, dann ergibt das die vorgesehenen vier für den Proposer, einen für den ersten Responder, einen weiteren für den zweiten Responder usw. Im Falle von zwei Respondern also 4 + 1 + 1 = 6 anstatt 4 + 0.5 + 0.5 = 5. Im Falle von vier simultan entscheidenden Respondern weitet

[26] Die ursprünglichen und bearbeiteten Fassungen finden sich im Anhang ab Seite 89.

[27] „Bitte beantworten Sie folgende Frage: Ein Ball und ein Schläger kosten zusammen 1.10°Euro. Der Schläger kostet 1.00°Euro mehr als der Ball. Wie viel kostet der Ball?“ Die meisten Personen antworten mit 0.10°Euro, korrekt wären allerdings 0.05°Euro.

sich dieser Betrag entsprechend aus. Wie unter Abb. 4 bereits erwähnt wurde, ist dies eigentlich unmöglich und vor allem problematisch, wenn mindestens zwei Responder das Angebot akzeptieren, weil dadurch Mehrkosten entstehen. Trotzdem wurde dieses Verfahren beibehalten, aus mehreren Gründen. Zum einen soll die Spielsituation nicht zu sehr vom Original aus Bern abweichen. Das Risiko ist gegeben, dass die Responder so etwas durchschauen, es wird aber nicht darauf abgezielt und so gut wie möglich verschleiert. Zum anderen sollen bei konstantem Anteil für den Proposer (vier von fünf Euro) die variablen Anteile für die vier Responder nicht zu klein sein (0.25°Euro), was das altruistic punishment noch kostengünstiger werden ließe. Weiterhin würde bei fixierten Anteilen der vier Responder (z. B. 0.50°Euro) der variable Anteil des Proposers in der Perspektive mancher Responder vielleicht nicht ausreichend unfair erscheinen. Ein eher pragmatischer Grund war letztlich, dass somit die wenigen vorgenommenen Änderungen die Programmierung des Experimentes nicht unnötig komplizierter machten.
Der im Anschluss an das Experiment obligatorische Fragebogen wurde nicht in seiner originären Version aus Bern übernommen. Der in der Reproduktion verwendete Fragebogen basiert auf der standardisierten Vorlage des „LEx“ und wurde nur geringfügig abgeändert[28]. Denn der Fragebogen in Bern war zu umfangreich und umfasste zahlreiche, inhaltlich weniger bedeutsame Items (Vertrauen in Institutionen, Personen, Netzwerkdichte etc.). Außerdem lag zu Zwecken der Reproduktion nur in Form des Codebuchs vor und war nicht bereits in z-tree implementiert.

6.2 Durchführung

Zu Beginn des Wintersemesters 2014/2015 wurden die Teilnehmer aus dem Probandenpool des Leipziger Experimentallabors (LEx) ausgewählt und eingeladen. Dabei wurden ebenso wie in Bern neben der genauen Anschrift des Labors noch Informationen zur Dauer des Experiments, dem durchschnittlichen Verdienst und der Aufwandsentschädigung mitgeteilt. Letztere lag zunächst bei drei Euro, wurde aber wegen mangelnder Beteiligung umgehend auf fünf Euro erhöht. Für einen etwaigen daraus resultierenden Methodeneffekt wird in der statistischen Analyse kontrolliert.
Die Experimente wurden in der Zeit vom 24. Oktober bis 14. November 2014 im PC-Pool des Geisteswissenschaftlichen Zentrums durchgeführt.

[28] Es wurden ein paar Fragen der „Self Report Altruism Scale“ (Rushton et al. 1981) hinzugefügt. Ein höherer Summenwert („score“) indiziert mehr berichtete altruistische Handlungen außerhalb des Labors. Im Anhang befindet sich die Skala im Original und mit den ausgewählten Items in (eigenverantwortlicher) Übersetzung ins Deutsche.

Nach der Erläuterung des Ablaufs und der Anweisung durch den Autor der Arbeit konnten Fragen gestellt werden. Sodann wurden die Spiele begonnen. Nach Beendigung des Fragebogens wurden die Probanden einzeln aufgerufen, sich ihre erspielte Auszahlung von den Laborleitern im vorderen Bereich des PC-Pools abzuholen und dies mit einer Unterschrift zu bestätigen. Somit entstanden zwar keine Warteschlangen an der Auszahlungsstelle, trotzdem war diese Situation nicht so aufwendig anonym gestaltet beziehungsweise durch einen Nachbarraum „doppelblind“, wie dies in Bern der Fall war.

7 Ergebnisse

7.1.1 Univariat

An der Reproduktion nahmen insgesamt 115 Personen teil. Die Stichprobe der Versuchspersonen unterscheidet sich mich 23.7 Jahren und einem Frauenanteil von 70 % nur unwesentlich von der in Bern. Es befanden sich hauptsächlich junge Studenten mit im Mittel 4.5 absolvierten Semestern darunter. Die wenigsten davon hatten einen Studienhintergrund nahe der Spieltheorie oder Logik, so dass nur 19 % beziehungsweise 23.5 % bereits ein entsprechendes Buch gelesen oder an einem solchen Lehrangebot teilgenommen hatten. Allerdings gab es nur sieben Studenten der Wirtschaftswissenschaften, weshalb diese in der weiteren Analyse nicht berücksichtigt werden.
Mit 51 Personen spielten etwa 44% das Teildesign mit zwei, der Rest das mit vier Respondern. Das fingierte Angebot von 4:1 beziehungsweise 4.5:0.5 wurde gleich häufig gestellt und insgesamt lehnten 59 % der Probanden diese Aufteilung ab. Dieser Wert liegt erneut höher als erwartet und Hypothese 3 kann nicht verifiziert werden. Im Gegenzug wurden durchschnittlich 2.16°Euro von 5°Euro angeboten, was als ausreichend fair gelten kann und sogar noch etwas näher an der 50-50-Marke liegt als der Wert 3.84 von zehn Franken in Bern. Der neben der Aufwandsentschädigung erspielte Verdienst lag im Mittel bei 3.08°Euro, von maximal möglichen 5.50°Euro. Die Teilnehmer konnten das Experiment wie in der Einladung angekündigt mit durchschnittlich 7.57°Euro (Median =7.50) verlassen. Die Probanden berichteten im Fragebogen ein minimal akzeptables Angebot (folglich MAO) von durchschnittlich 1.47°Euro (Median =1.50, Werte von 0.1 bis 4.9). Dieser Befund ähnelt dem in Bern mit 3.20 von zehn Franken (Median=4, Werte von 0 bis 5).
Die Probanden gaben sowohl ein gutes oder sehr gutes Verständnis als auch Anonymitätsempfinden des Experiments an. Obschon eigentlich mit derartigen Experimenten unerfahrene Studenten die Zielgruppe bildeten und in der Stichprobe hätten auftauchen sollen, war dennoch fast ein Viertel damit vertraut und 5 % nicht-studentisch. Die zwischen die eigentlichen Tauschsituationen eingestreute Quizfrage beantworteten lediglich 37 % korrekt, die meisten gaben intuitiv „10 Cent“ an.

7.1.2 Bivariat

Die Ablehnungsraten unterscheiden sich für beide Treatments nur geringfügig und überdies nicht signifikant (57 % beziehungsweise 61 % im vier-Responder-Fall, t=-0.44). Trotz der erneut recht hohen Ablehnungsraten ist genügend Verhaltensvarianz vorhanden, die Probanden agieren nicht ausschließlich fair, sondern auch egoistisch. Die absoluten Werte liegen zwar leicht unter denen in Bern, aber der antizipierte Effekt der Verantwortungsdiffusion, d.h. geringere Ablehnungsraten im

zwei-Responder-Fall und ebenfalls geringere als im normalen Ultimatumspiel, stellt sich anscheinend auch hier nicht ein. Er liegt zwar etwa 10 % unter den Werten in Bern. Trotzdem gibt es hier keine Evidenz für die theoretischen Überlegungen.

Das fingierte Angebot von 4:0.5 Euro wurde zu 69 % abgelehnt und damit knapp signifikant häufiger als das von 4.5:1 Euro (49 %, t=2.19). Der erste Wert gleicht sich dem in Bern an, letzterer ist kompatibel mit dem Forschungsstand zum normalen Ultimatumspiel. Hier deutet sich eine mit der Theorie konforme Tendenz an, wonach das billigere altruistic punishment häufiger ausgeführt wird.

Die während der Erhebungsphase erhöhte Teilnahmevergütung von drei auf fünf Euro besaß keinen Effekt auf die Ablehnungsraten und auch nicht auf die Höhe der eigenen Abgaben.

Testpersonen, die die Vorleistung des fiktiven Proposers annahmen, erhielten „natürlich“ signifikant mehr als die Vergleichspersonen. Allerdings liegt der Wert mit durchschnittlich 1.10 Euro höher als erwartet und ist nicht alleinig über die 4:1-Aufteilungen erklärbar. Offensichtlich haben diese Personen weiterhin clever gespielt.

Die eigenen Vorschläge der ablehnenden Probanden liegen mit 2.29°Euro signifikant um 32 Cent über denen der annehmenden Personen (t=-2.45). Dieses Ergebnis gilt auch aufgegliedert nach Treatments (nicht signifikant im 4-Respoder-Fall) und deckt sich mit dem Befund aus dem Experiment in Bern. Selbige verlangten laut Fragebogen etwa 1.05 Euro mehr als akzeptierende Personen (MAO für letztere im Mittel bei 0.85 Euro, t=-7.99). Fairness scheint demnach ein Anspruch an andere und sich selbst zu sein, es wird nicht indirekt reziprok gehandelt, obschon der Effekt etwas schwächer als in Bern ausfällt.

Trotz der recht fairen eigenen Offerten, traten vereinzelt auch hyperfaire Angebote auf, die ein schlechtes Verständnis nahelegten. Die Prüfung ergab einen stark negativen signifikanten Zusammenhang für Angebote jenseits der 50-50-Marke (r=-0.74***, n=13) und wurde für das Design mit vier Respondern noch gesteigert (r=-0.82**, n=7). Einige wenige Probanden haben das Experiment also tatsächlich falsch verstanden, orientierten sich dennoch an der fairen Aufteilung.

Der erspielte Verdienst (Profit) als Stellvertreter für sehr „rationales“ oder weniger faires Verhalten fällt geringer aus, je mehr jemand als MAO verlangt (r=-0.50***) und je höher das Semester ist (r=-0.29***), die Semesterzahl hängt ähnlich stark mit dem MAO zusammen (r=0.23**). Höhere Semester sind also fairer, fairere Personen verlangen fairere Angebote und geben diese selbst auch vor. Frauen lehnten das empfangene Angebot zu 60.5 % ab, Männer zu knapp 56 %, womit sich keine gewichtigen oder signifikanten Unterschiede ergeben.

Die Variable Anonymität besitzt keine besonderen Einflüsse auf das Spielverhalten und ist damit ähnlich unbedeutend wie in Bern, dasselbe gilt für die Erfahrung mit Experimenten. Anscheinend können beide Faktoren keine egoistischeren Verhaltensweisen befördern.

7.1.3 Multivariat

Alleinig der Effekt der Verdopplung der Responderzahl von zwei auf vier erhöht die vorhergesagte Chance auf Ablehnung des fingierten Angebots nicht signifikant um den Faktor 1.18. Der Effekt ist bis dato relativ gering und auch entgegen der bisherigen Beobachtungen zum bystander effect. Die Hypothesen 1 und 2 können an dieser Stelle nicht gestützt werden.

Im Modell, das ausschließlich die Höhe der eigenen Abgaben beinhaltet, erhöht sich mit jedem weiteren abgegebenen Euro die vorgesagte Chance auf Ablehnung signifikant um den Faktor 2.01. Sofern beide Haupteffekte zusammen aufgenommen werden, liegen die Odds-Ratios für vier simultane Responder bei 1.32 und für die eigenen Vorschläge bei 2.06**. Die Wirkung der Design-Variable ist in der Grundgesamtheit wahrscheinlich null und verbessert die Schätzung in diesem dritten Modell nicht.

Im vierten Modell wird die zu diesen beiden zentralen Effekten die Höhe des fingierten Angebots hinzugezogen, die selbst die geschätzte Chance auf Ablehnung mit dem Faktor 0.33*** verringert. Sie lässt die eigene Vorschlagshöhe auf 2.53*** ansteigen, ändert praktisch nichts an der Wirkung des Treatments und ist essentiell für die Modellschätzungen.

Die Höhe der Aufwandsentschädigung in Modell 5 hat bei einem Odds-Ratio-Wert von 0.87 einen vernachlässigbar kleinen negativen Einfluss auf die Ablehnung des Angebots. Sie modifiziert die Werte der anderen drei Variablen praktisch nicht und scheint an diesem Punkt weniger wichtig für die Schätzungen. Das Modell 6 ohne die Höhe des fingierten Angebots und mit einer Interaktionsvariable aus Teilnahmevergütung und Design-Variable steigert hingegen den Odds-Ratio-Wert des Designs schlagartig auf 6.01**, senkt den der eigenen Abgaben minimal auf 2.38*** und erhöht das der Aufwandsentschädigung selbst auf 2.11. Der entsprechende Wert dieser Interaktionsvariable selbst liegt bei vergleichsweise schwachen 0.14*. Probanden im 4- Responder-Design mit der höheren Teilnahmevergütung von fünf Euro haben damit eine etwas geringere geschätzte Chance der Ablehnung als ihre Mitspieler (Probanden im 2-Responder-Design mit garantierten zwei oder vier Euro und Probanden im 4-Responder-Design mit garantierten zwei Euro). Das altruistic punishment stellt sich damit etwas weniger wahrscheinlich ein, wenn vier simultan entscheidende Probanden mehr Geld zugesichert bekamen – wenn aus Egos Sicht also genügend andere Responder die vergleichsweise günstige Sanktion ausüben könnten (im Einklang mit den Vorüberlegungen). Allerdings ist der Wert überlagert von den deutlich stärkeren anderen Hauptvariablen mit entgegengesetzter Wirkrichtung. Der Effekt dieser Interaktionsvariablen ist dennoch in der Grundgesamtheit sehr wahrscheinlich nicht null.

Wird nun in Modell 7 wieder die Höhe des künstlichen Proposer-Angebots integriert, so nimmt diese ihren vorigen Wert von 0.32*** an und ändert nur die Werte der Design-Variable von 6.01** auf 6.90** und den der eigenen Vorschlagshöhe von 2.38*** auf 2.99***. Dieses Modell ist besser als seine beiden Vorgänger Modell 5 und 6.

In Modell 8 wird die Interaktion aus dem Design und dem fingierten Angebot aufgenommen. Damit haben Probanden im 4-Responder-Design und künstlich angebotenen 1 von 4 Euro eine um den Faktor 0.62 geringere vorhergesagte Chance auf Ablehnung als ihre Mitspieler (also Probanden im 2-Responder-Design mit Angeboten von 4.5:0.5 oder 4:1 und Probanden im 4-Responder-Design mit dem Angebot von 4.5:0.5). Die Verantwortungsdiffusion hinsichtlich des sanctioning dilemma stellt sich damit etwas wahrscheinlicher ein, wenn das altruistic punishment für die Probanden teurer wird, also gemäß der Vermutungen unter Hypothese 4. Aber auch dieser Effekt ist praktisch bedeutungslos im Vergleich mit den wesentlich stärkeren Faktoren des Treatments von 9.20** und der eigenen Vorschlagshöhe von 3.08*** und ist gemäß LR-Test in der Grundgesamtheit wahrscheinlich null.

In einem letzten Modell wird schließlich eine Interaktion aus der Teilnahmevergütung und dem fingierten Angebot hinzugefügt. Diese fällt mit einem Odds-Ratio-Wert von 0.67 fast identisch aus wie die Interaktion des Vorgängermodells und verringert die vorhergesagte Chance auf Ablehnung. Die Probanden mit einem erhöhten showup-fee von fünf Euro, die das Angebot von 4:1 erhielten, tendieren zu weniger Ablehnungen als ihre Mitspieler, denen eine Aufteilung von 4.5:0.5 Euro angeboten wurde, oder die drei Euro showup-fee erhielten. Ein relativ teureres altruistic punishment wird also seltener erteilt, was ähnlich den Modellen 6 beziehungsweise 8 positive Evidenzen für die theoretischen Annahmen liefert. Dennoch wird auch dieser Interaktionseffekt von den wesentlich stärkeren Wirkungen der Design-Variable (8.75**) und der eigenen Abgabenhöhe (3.09***) überdeckt, mithin bringt er keine Verbesserung der Vorhersage. Selbst der bloße Effekt der höheren Aufwandsentschädigung von fünf Euro ist in den Modellen 5 bis 8 mindestens doppelt so stark und steigert die vorhergesagte Chance der Ablehnung, wenn auch nicht signifikant. Die Befürchtung, dass eine höhere Aufwandsentschädigung zu riskanterem oder egoistischerem Verhalten verleitet, kann auch multivariat nicht bestätigt werden.

Dieses Modell mit sämtlichen Interaktionsvariablen wurde analog als LP-Modell gerechnet, die Effekte weisen dabei in dieselben Richtungen (inklusive ihrer Signifikanz). Deren VIF-Werte liegen allesamt unter 6 und sind daher unbedenklich. In einem hier nicht präsentierten Modell wurden Kontrollvariablen der Soziodemographie integriert, die praktisch keine Änderungen an den Zusammenhängen hervorrufen und selbst keine signifikanten Effekte produzieren.

Auch in der multivariaten Analyse der Reproduktion zeigen sich kontraintuitive Tendenzen, welche die Hypothesen im Wesentlichen nicht stützen können, obschon in den letzten Modellen Ansätze

sichtbar wurden. Es zeigt sich ein den Berner Modellen ähnliches Schema, indem die zentralen Variablen in dieselben Richtungen deuten, die Modelle mit den wichtigen Haupteffekten bereits gute Vorhersagen leisten und die Erweiterungen keine besonderen Verbesserungen der Varianzaufklärung erzielen[29]. Die Resultate verifizieren keine der Hypothesen.

Im unmittelbaren Vergleich des Ablehnungsverhaltens zweier simultan agierender Responder in Bern und Leipzig zeigt sich, dass die Höhe der eigenen Abgaben die geschätzte Chance der Ablehnung des fingierten Angebots steigert. In Bern geschieht dies um den stabilen Faktor 1.59 (nicht signifikant) und in Leipzig schwankt der Wert je nach zusätzlichen Regressoren zwischen 4.66** und 6.55**. Die Höhe des künstlichen Angebots fällt in Bern mit 0.67 beziehungsweise mit 0.34* in der Reproduktion ähnlich aus wie in den obigen Modellen. Die Effekte deuten in beiden Experimenten in gleiche Richtungen, sind in Leipzig aber nahezu doppelt so stark und mithin signifikant.

[29] Eine Interaktionsvariable aus den Haupteffekten des Treatments und der Vorschläge ist nicht mit im Modell aufgenommen, da sie ihre VIF-Werte im Anschluss an ein LPM jeweils bei etwa 19 lagen (in einem Modell nur mit Vorschlag, Design und Interaktionseffekt: 14.9 für Design und 17.4 für Interaktions, Vorschlag nur 4.6). Andere Interaktionsvariablen hingegen haben wesentlich geringere Werte, und zwar höchstens 5.80 und sind damit unauffällig.

Die Pearson-Residuen der Regressionsmodelle sind ähnlich wie in Bern gemäß Shapiro-Francia- und Shapiro-Wilk-Test nicht normalverteilt. Die meisten weisen eine Ballung um den Wert 0 auf, einige wenige Abweichungen verlaufen bis (-6) oder +3. Für alle existieren jeweils einige einflussreicher Fälle, für die die Zusammenhänge offenbar schlicht nicht zutreffen. Aus selbigen Gründen wie oben erwähnt werden sie von der Analyse dennoch nicht ausgeschlossen.

Abhängige Variable: Ablehnungsverhalten, Leipzig 2014

Unabhängige Variablen	Modell 1	Modell 2	Modell 3	Modell 4	Modell 5	Modell 6	Modell 7	Modell 8	Modell 9
Haupteffekte									
Design	1.18		1.32	1.39	1.41	6.01**	6.90**	9.20**	8.75**
eigene Angebote		2.01**	2.06**	2.53***	2.55***	2.38***	2.99***	3.08***	3.09***
fingiertes Angebot				0.33***	0.33***		0.32***	0.40	0.56
showup-fee					0.87	2.11	2.12	2.13	2.76
Interaktionen									
Design & showup-fee						0.14*	0.12**	0.12**	0.12**
fingiertes Angebot & Design								0.62	0.67
fingiertes Angebot & showup-fee									0.61
Konstante	1.31	0.32*	0.27**	0.30***	0.32**	0.12**	0.13**	0.10**	0.09**
Pseudo- R^2	0.00	0.04	0.04	0.09	0.09	0.07	0.12	0.12	0.12
BIC	-380.8	-386.7	-382.4	-385	-380.3	-377	-380	-375	-371
(adjustiertes) Count-R^2	0.59 (0.00)	0.64 (0.11)	0.63 (0.11)	0.64 (0.11)	0.64 (0.13)	0.64 (0.13)	0.67 (0.19)	0.68 (0.21)	0.68 (0.21)

Abb. 6: Ablehnungsverhalten der Responder in Leipzig.

Logistische Regression auf die Entscheidung, das unfaire Angebot abzulehnen (abhängige Variable, Wert 0=akzeptiert, 1=nicht akzeptiert). Angabe der Koeffizienten in Odds-Ratios. Signifikanzniveaus: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Das Pseudo-R^2 ist das von McFadden. Für alle Modelle liegt das N bei 115. Kodierung Design: Wert 0= 2 Responder, 1= 4 Responder (n=64); fingiertes Angebot: Wert 0= 4.50 zu 0.50°Euro, 1= 4 zu 1°Euro; showup-fee: Wert 0= 3°Euro, 1= 5°Euro (n=86).

Abhängige Variable: Ablehnungsverhalten, Bern 2013 und Leipzig 2014

Unabhängige Variablen	Bern 1	Bern 2	LEx 1	LEx 2	LEx 3
Haupteffekte					
eigene Abgaben	1.58	1.59	4.66**	5.64**	6.55**
fingiertes Angebot		0.67		0.34*	0.35
showup-fee					2.47
Interaktion					
fingiertes Angebot & showup-fee					
Konstante	0.52	0.61	0.04**	0.05*	0.02**
N	39	39	51	51	51
Pseudo- R^2	0.04	0.04	0.09	0.13	0.15
BIC	-92	-89	-128	-128	-125
(adjustiertes) Count-R^2	0.77 (0.10)	0.77 (0.10)	0.65 (0.18)	0.67 (0.23)	0.69 (0.27)

Abb. 7: Vergleich des Ablehnungsverhaltens zwei simultan agierender Responder in Bern und Leipzig.

Logistische Regression auf die Entscheidung, das unfaire Angebot abzulehnen (abhängige Variable, Wert 0= akzeptiert, 1= nicht akzeptiert). Angabe der Koeffizienten in Odds-Ratios. Signifikanzniveaus: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Das Pseudo-R^2 ist das von McFadden.

8 Methodendiskussion

Bei näherer Untersuchung traten auch in diesem Experiment Probleme auf, die teilweise dem Pragmatismus einer „einfachen“ Reproduktion geschuldet sind. So gabe es keine echte Kontrollgruppe, sondern diejenige in Bern wurde stellvertretend als solche zu Rate gezogen. Deren Ergebnisse lassen sich einigermaßen gut an die zu Beginn dargelegten Forschungsergebnisse anfügen. Trotzdem bleibt die naive Annahme, dass dabei mangels Komplexität weniger schief lief als im Treatment mit simultanen Respondern.

Die eigentlich wichtigen Informationen über die Rundenzahl, das Re-Arrangement oder Neubesetzung der Spielpositionen waren sowohl in Bern als auch in Leipzig äußerst spärlich beziehungsweise fehlten gänzlich – anders als in bisherigen Studien üblich. De Facto spielten die Probanden alle zunächst gegen fiktive Proposer, danach gegen künstliche Responder. Sie wussten aber nicht, ob sie mit ihrer eigenen Offerte den soeben unfairen Proposer bestrafen könnten, also direkt reziprok handeln oder ob sie damit irgendeinen Mitspieler im Raum getroffen hätten, der dann zweimal unfair behandelt worden wäre. Diese Problematik wird noch verstärkt, indem es sowohl in Bern als auch in Leipzig keinerlei Übungsrunden ohne Effekt auf die Auszahlungen gab. Die Probanden hatten nicht die Möglichkeit, etwaige Fehlentscheidungen in Folgerunden auszubessern oder vor der eigentlichen Tauschsituation Verständnisfragen zu stellen. Der Vorteil daran sind die oben genannten unterbundene Lerneffekte, trotzdem ist nicht gewiss, ob die Tauschsituation echt verstanden wurde (beziehungsweise inwieweit Zweifel und Unsicherheit das „echte“ Verhalten verzerrten). Nach Binmore (1998, S. 30) sind die Vorhersagen der Rational-Choice-Theorie annähernd deckungsgleich mit den Daten, wenn die Entscheidungssituation leicht verständlich, die Lernphase ausreichend ist und adäquate Anreize für die Versuchspersonen vorliegen. Die hier präsentierten Diskrepanzen zwischen Theorie und Empirie sind zumindest teilweise auf das Fehlen dieser Faktoren zurückzuführen.

Es wurde in der Reproduktion zwar nur eine Verständnisfrage seitens der Probanden während der Tauschsituationen gestellt, aber diese war eher einfacher Natur. In der Reproduktion hat auch nur eine Person durchschaut, dass der „Kuchen“ größer wird und sich die Werte bei Akzeptanz aller Responder auf mehr als 5 Euro aufsummieren. Zumindest war sie die einzige, die dazu im Nachgang des Experiments mit den Laborleitern ins Gespräch trat. Spekulativ bleibt, ob andere Probanden dieselben Probleme hatten, sich aber nicht trauten, nachzufragen.

Leider kann nicht beurteilt werden, inwiefern die gewollte Täuschung der Teilnehmer der Reproduktion mittels der eingestreuten Quizfrage auch tatsächlich erreicht wurde. Vielleicht waren sie im Übermaß auf die korrekte Beantwortung konzentriert, so dass ein nicht zusätzlich unintendierter, aber auch nicht überprüfbarer Methodeneffekt vorliegt.

Ein anderes Problem bei den echten Angeboten der Studienteilnehmer ist das Fehlen der Variable darüber, ob diese Angebote angenommen oder abgelehnt wurden. Es kam zwar auf diese fairen oder eben unfairen „echten“ Abgaben an, aber die Tauschsituation ist damit eigentlich nicht konsequent zu Ende geführt worden.

Die Erweiterung des Standard-Fragebogens des LEx um Elemente der Altruismus-Skala brachte in der Analyse nicht besonders fruchtbare Erkenntnis und hielt die Probanden länger im Labor als unbedingt notwendig.

Die multivariate Analyse hätte sich intensiver mit den mehr oder minder einflussreichen Fällen befassen können, was hier in Anbetracht des Umfangs der Arbeit nicht verfolgt wurde. Es ist möglich, dass Ausreißer oder Extremwerte Artefakte produzieren. Ein testweiser Ausschluss dieser verbesserte den Bestätigungsgrad der Hypothesen vielleicht.

Trotz der zahlreichen Nachteile beim Experiment in Bern könnte dieser Datensatz mit dem der Reproduktion zu einer Datei zusammengefügt werden – mit einer entsprechenden Dummy-Variable für den Erhebungsort und der Anpassung der Geldbeträge. Danach hätte die Analyse den Vorteil einer größeren Fallzahl, sogar unter Ausschluss äußerst einflussreicher Fälle und auch im Hinblick auf die in Bern vorhandene Kontrollgruppe zum normalen Ultimatumspiel ohne multiple Responder.[30]

[30] Mithin könnten dann andere Regressionsverfahren zum Zuge kommen, so z.B. Mehrebenenmodelle, bei dem die Individuen als in der Gruppe genisteter Faktor behandelt werden, oder auch Propensity-Score-Matching, bei dem jedem Fall passende Partner in der jeweils anderen Gruppe (mit beziehungsweise ohne multiple Responder) gesucht werden. Varriierende Verfahren mit ähnlichen Ergebnissen könnten die Befunde zusätzlich erhärten.

9 Schlussbetrachtungen

Der Datenlage nach sorgten multiple Responder in beiden Experimenten nicht dafür, dass sich der sonst empirisch gut beobachtbare und bestätigte Effekt der Verantwortungsdiffusion erneut deutlich zutage tritt. Er konnte sich multivariat gegenüber anderen Einflüssen nicht behaupten, sondern fiel schwächer aus und wies außerdem (signifikant) in die Gegenrichtung. Die hier präsentierten Hypothesen konnten nicht untermauert werden. Dies kann mehrere Gründe haben:

Kulturelle Unterschiede zwischen Studenten in der Schweiz und Deutschland sind natürlich nicht von der Hand zu weisen und sorgen sicherlich für etwas Verhaltensvarianz. Sie sind jedoch vermutlich vernachlässigbar, vor allem im Hinblick auf stärkere, bessere erforschte Unterschiede zwischen den typischen Versuchspersonen im Labor[31] und indigenen Stammesgesellschaften (Henrich et al. 2001).

Die Ergebnisse lassen einerseits die Schlussfolgerung zu, dass eine altruistische Sanktionierung aus einem emotionalen Reflex oder der wütenden überkompensierenden Reaktion erzürnter Akteure über derart unfaire Angebote resultiert (Camerer 2003, S. 49; Nikiforakis 2013, S. 211f.), obgleich das nicht genauer ergründet wurde mittels Fragebogen. Andererseits ist das altruistic punishment wahrscheinlich ein low-cost-Verhalten, das nur bei entsprechend niedrigen Kosten ausgeübt wird. Bei der Reproduktion waren die Kosten dafür (0.50 beziehungsweise 1°Euro) im Verhältnis zum Profit (5°Euro sicherer showup-fee und zweimal 5°Euro zu verhandeln) zwar etwas höher als die 1 oder 2°Franken „teure“ Sanktion in Bern (20°Franken fix und zweimal 10°Franken zur Disposition), aber dennoch eher zu niedrig.

Die uneigennützige Sanktion unterliegt, wie oben erwähnt, zwei häufig verletzten moralischen Schranken: 1) Wer im Glashaus sitzt, sollte nicht mit Steinen werfen. 2) Man sollte eigens erfahrenen Ärger nicht an Dritten auslassen (Elster 2005, S. 206). In dem Experiment betraf dies genau die Übereinstimmung der Reaktion auf das künstliche unfaire Angebot und die anschließenden eigenen Abgaben. Offenbar haben die Akteure diese beiden Leitmaximen nicht systematisch verletzt und nicht indirekt negativ reziprok gehandelt, indem sie zwar empört das Angebot ausschlugen, aber selbst kein besseres Beispiel darstellten und aggressive Offerten anboten. Dass faire Personen fair spielen (weil sie es mögen) generiert keine Erkenntnis und wäre mithin tautologisch – vor allem bei einer operationalen Definition von Fairness: Fair ist, wer fair spielt. Dennoch passiert das anscheinend unabhängig vom hiesigen experimentellen Treatment. Die Probanden verhalten sich in allen Experimentalgruppen recht fair, nicht reziprok und bestrafen unfaires Verhalten. Dies steht

[31] Sogenannte „WEIRD“- People: white, educated, industrialized, rich, democratic – eben Studenten an Hochschulen und Universitäten westlicher Industriegesellschaften.

im Einklang mit anderer empirischer Evidenz, so z. B. Grosskopf (2003), die keine signifikanten Unterschiede feststellen kann zwischen Probanden mit Rollenwechsel und Nicht-Rollenwechsel hinsichtlich ihrer eingenommenen Spielerpositionen als Proposer oder Responder. Fraglich ist allerdings, ob die Probanden hier rachsüchtig agiert hätten, wenn sie den ihnen gegenüber zuvor unfairen Proposer in einer zweiten Interaktion wiedergetroffen hätten. Eine Aussage über direkte negative Reziprozität und ihre Stärke im Vergleich zur Fairness (Diekmann 2004, S. 502) ist damit leider nicht möglich.

Abgesehen davon ist nicht klar, ob aggressive Angebote aus Angst vor counter-punishment abgelehnt wurden oder ob faire Angebote aus Angst vor altruistic punishment der anderen Responder erteilt wurden, also strategisch motiviert und sozial erwünscht gehandelt wurde. Daran schließt sich an, dass sie vermutlich auf weitere Spielrunden gehofft hatten, um den Proposer zu besseren Angeboten zu bewegen („educative punishment“), was letztlich ausblieb.

Jedenfalls sind die Daten ein Indiz, wonach Fairness eine über Sozialisation internalisierte Präferenz ist anstatt eine durch Sanktion gestützte soziale Norm (Bicchieri und Xiao 2009, S. 192), denn die Anonymität hat nicht zu egoistischerem Verhalten verleitet.

In dem zentralen Experiment kann trotz entsprechender Gestaltung als symmetrisches Freiwilligendilemma (alle Versuchspersonen mit demselben showup-fee und ohne weitere Anfangsausstattung versehen) ein asymmetrisches Freiwilligendilemma vorgelegen haben. Wie oben bereits erwähnt, fällt die Kooperation dabei für gewöhnlich höher aus und der starke Spieler legt eher sein Veto ein als schwächere Spieler. Das wäre eine Erklärung dafür, warum relativ wenig Trittbrettfahrer vorhanden sind. Leider ist nicht weiter überprüfbar, wie hoch der Nutzen (ein möglicher „warm glow“) individuell ausfällt. Das Einkommen für eine Bewertung der Kosten des altruistic punishment hinzuzuziehen, ist aufgrund geringer Varianz in der Population der Studenten nicht sonderlich aufschlussreich, denn die Sanktion bleibt mit 0.50 oder 1°Euro ähnlich „teuer“ – egal ob jemand 300 oder 750°Euro (beziehungsweise Franken) monatlich zur Verfügung hat (das gilt umso mehr für Schweizer Finanzverhältnisse)[32]. Eine Überprüfung auf anderer Datenbasis mit einer speziellen Population oder (relativ) höheren Spielbeträgen bietet sich an.

Die Daten des hiesigen Experiments könnten in einem nachfolgenden Schritt mit Theorien abgeglichen werden, die nebst egoistischen Präferenzen auch solche in Bezug zu anderen Mitspielern in ihren auszahlungsbasierten Modellen berücksichtigen (Güth und Kocher 2014). Wegen des Umfangs konnte dies in der vorliegenden Arbeit leider nicht verfolgt werden. Aber auch an diesem Punkt ergäben sich eventuell Schwierigkeiten, für die das Design verantwortlich ist. Es besteht unter den Probanden vielleicht Unklarheit über den genauen Referenzpunkt oder die

[32] Eventuell bringt die Frage Aufschluss, ob jemand (staatliche) Studienbeihilfe erhält oder nicht.

Referenzgruppe (Ockenfels 1999, S. 187) und zwar wegen eines etwaigen Missverständnisses der Tauschlogik: Die falsche Schlussfolgerung mancher Probanden war vielleicht die Aufsummierung der Proposeranteile, welche die empfundene Unfairness noch vergrößert hätte (vier Angebote zu je 4:1 Euro ergibt – gegebenenfalls alle Responder akzeptieren – 16°Euro für den Proposer). Womöglich war auch das der individuelle Richtwert. Die den Designs zugrundeliegende Idee war so oder so anscheinend noch nicht ausgereift genug.

Die Ergebnisse deuten am Rande auf die Bedeutung der Entscheidungszeit als zukünftig zu untersuchenden Faktor. Und zwar nicht nur, weil das altruistic punishment zumindest bei manchen Probanden sehr spontan und emotional aufgeladen erfolgt. Wenn die Versuchspersonen sich bei ihren eigenen Angeboten schneller entschieden, dann fielen diese Vorschläge auch etwas höher aus (r=0.18*). Vielleicht wird die Fairnessnorm weniger in Zweifel gezogen, sondern umgehend präsentiert. Außerdem reagierten ältere Probanden und solche in höheren Semestern, Probanden mit einem höheren Anonymitätsempfinden langsamer auf das künstliche Angebot des Proposers (r=-0.23**, r=-0.1935**, r=-0.18*). Die ersten beiden erscheinen vielleicht etwas bedachter, womöglich etwas verwundert über die Aggressivität des Angebotes, letztere ziehen unter den anonymen Umständen vielleicht ein egoistischeres Verhalten in Betracht[33]. Personen mit Experimentalerfahrung beantworteten hingegen 3.33 Sekunden schneller (t=-1.81) und impulsivere Reaktionen dieses Typs stehen in schwachem positiven Zusammenhang mit zügigen Antworten auf die Quizfrage (r=0.18*) und negativ mit der vermuteten Anzahl ablehnender Responder (r=-0.17*)[34].

Den meisten Laborexperimenten kann wohl folgender Vorwurf gemacht werden: „[...They...] are lousy *simulations* of the social dilemmas with which most of us are concerned." (Dawes 1980, S. 188; Hervorhebung im Original). Gemäß Güth und Kocher (2014) gibt es praktisch kaum Untersuchungen, die das Verhalten im Labor mit dem im Feld verknüpfen. Generell sind Feldexperimente zu Verhandlungssituationen eher die Ausnahme (Güth und Kocher 2014, S. 10). Hierin läge ein durchaus schwierig umzusetzender, aber interessanter Ansatzpunkt für weitere Forschung. Benz und Meier (2008) bilanzieren einen relativ starken signifikanten Zusammenhang

[33] Nach subjektivem Eindruck der Laborleiter ließ sich in puncto Anonymität ein Methodeneffekt feststellen. Demnach verhielten sich Probanden am Rand des PC-Pools und insbesondere ohne unmittelbaren Nachbarn egoistischer. Das könnte zukünftig systematisch berücksichtigt werden.

[34] Dr. Andreas Tutić und Sascha Grehl führten am hiesigen Institut für Soziologie Experimente zum „iterated reasoning" durch. Es geht unter anderem hervor, dass auf zügige Antworten reflexive Fragen häufiger falsch sind und Frauen eher zu diesen tendieren. Hier konnte sich anhand der Quizfrage (Baseballschläger) bestätigen, dass falsche Antworten um 27.5 Sekunden eher fallen (t=3.76), Frauen aber nur 8.25 Sekunden zeitiger (t=1.01) und nicht häufiger falsch antworten (63 %) als Männer (64.7 %).

zwischen Feld- und Laborverhalten der Studenten der Universität Zürich beim Entrichten eines Beitrags zu zwei Sozialfonds (Korrelationen variieren von 0.25 bis 0.4). In selbigem Forschungspraktikum in Bern wurde von einer Arbeitsgruppe ein Treatment vorgeschlagen, das mangels Ressourcen allerdings nicht durchgeführt werden konnte. Folgendes Szenario ist realistisch und könnte zu mehr externer Validität verhelfen: In einem Wartebereich vor dem Experimentallabor säßen zahlreiche echte Probanden und solche, die das nur vortäuschen. Ein offensichtlich gehbehinderter Schauspieler mit einem gebrochenen Fuß tritt heran und lässt eine seiner Krücken fallen. Die fingierten Probanden greifen nicht ein, denn sie erscheinen mit Zeitunglesen, Handyspielen oder Musikhören beschäftigt. Die tatsächlichen Probanden merken wahrscheinlich sodann, dass sie als vermeintlich Einzige vom Missgeschick erfahren haben. Mit dieser künstlich hergestellten Situation wird auf die Aufhebung des Freiwilligendilemmas abgezielt, indem die Probanden die Initiative ergreifen, selbstlos Hilfe anbieten und auch leisten. Als die abhängigen Variablen wären die Zeit bis zur Reaktion und ihre Stärke (bloßes Zuschauen, Nachfrage, Hilfestellung) messbar. Gut zu kontrollieren und zu variieren wären die Zahl der echten Probanden sowie Zahl der fingierten Probanden. Durch ein günstiges Arrangement könnte das Verhalten dieser echten Probanden in einem anschließenden experimentellen Design mit der Beobachtung abgeglichen werden (z. B. durch die Zuordnung der fingierten Probanden auf vorgegebene Sitzplätze). In einem anschließenden Fragebogen sollten neben den üblichen Fragen zur Sozialstatistik für einen zusätzlichen Abgleich der gesammelten Beobachtungen vor und im Labor ergründet werden, inwiefern eine Beteiligung der Probanden an gemeinnützigen Vereinen und Organisationen, Ehrenämtern, (z. B. Spendenbereitschaft zu humanitären Katastrophen oder für „open source“- und „crowd funding“-Projekte) vorliegt – so ähnlich, wie dies in der Reproduktion mittels Elementen der Altruismusskala versucht wurde.

Die Idee ist sicherlich unausgereift, erscheint bisweilen naiv und birgt jede Menge Schwierigkeiten. Trotzdem könnte man Daten gewinnen zum „echten“ Verhalten vor dem Labor, Handeln im Labor und dies mit berichtetem Verhalten zum Alltag abgleichen.

Literaturverzeichnis

Balafoutas, Loukas, und Nikos Nikiforakis. (2012). Norm enforcement in the city: A natural field experiment. *European Economic Review*, 56, 8:1773-1785.

Barclay, Pat, und Toko Kiyonari. (2014). Why Sanction? Functional Causes of Punishment and Reward. In *Social Dilemmas: New Perspectives on Reward and Punishment.* Hrsg. Paul van Lange, Bettina Rockenbach, und Toshio Yamagishi.

Barron, Greg, und Eldad Yechiam. (2002). Private e-mail requests and the diffusion of responsibility. *Computers in Human Behavior,* 18, 507–520.

Benz, Matthias, und Stephan Meier. (2008). Do people behave in experiments as in the field? - evidence from donations. *Experimental Economics,* 11, 286-281.

Bicchieri, Cristina, und Erte Xiao. (2009). Do the Right Thing: But Only if Others Do So. *Journal of Behavioral Decision Making*, 22, 191-208.

Binmore, Ken. (1998). Game Theory and the Social Contract, Volume 2: Just Playing. London: the MIT Press.

Boyd, Robert, Herbert Gintis, Samuel Bowles, und Peter J. Richerson. (2003). The evolution of altruistic punishment. *PNAS*, 100, 6:3531-3535.

Camerer, Colin F. (2003). Behavioral Game Theory: Experiments in Strategic Interaction. Princeton: Princeton University Press.

Cappelen, Alexander, Erik Sørensen, und Bertil Tungodden. (2010). Responsibility for what? Fairness and individual responsibility. *European Economic Review,* 54, 3: 429–441.

Charness, Gary. (2000). Responsibility and effort in an experimental labor market. *Journal of Economic Behavior und Organization*, 42, 375-384.

Croson, Rachel, und James Konow. (2009). Social Preferences and moral biases. *Journal of Economic Behavior und Organization*, 69, 3:201-212.

Cryder, Cynthia E., und George Loewenstein. (2012). Responsibility: The tie that binds. *Journal of Experimental Social Psychology*, 48, 441-445.

Darley, John, und Bibb Latané. (1968). Bystander intervention in emergencies: Diffusion of responsibility. Journal of Personality and Social Psychology, 8, 377-383.

Dawes, Robyn M. (1980). Social Dilemmas. *Annual Reviews Psychol*., 31, 169-193.

Denant-Boemont, Laurent, David Masclet, und Charles N. Noussair. (2007). Punishment, counterpunishment and sanction enforcement in a social dilemma experiment. *Economic Theory*,

33, 145-167.

Diekmann, Andreas. (1985). Volunteer's Dilemma. *The Journal of Conflict Resolution*, 29, 4:605-610.

Diekmann, Andreas. (1993). Cooperation in an Asymmetric Volunteer's Dilemma Game. Theory and Experimental Evidence. *International Journal of Game Theory,* 22, 75-85.

Diekmann, Andreas. (2004). The Power of Reciprocity. Fairness, Reciprocity, and Stakes in Variants of the Dictator Game. *Journal of Conflict Resolution*, 48, 487.

Diekmann, Andreas. (2008). Soziologie und Ökonomie: Der Beitrag experimenteller Wirtschaftsforschung zur Sozialtheorie. *Kölner Zeitschrift für Soziologie und Sozialpsychologie,* 60, 3:528-550.

Diekmann, Andreas. (2009). Spieltheorie: Einführung, Beispiele, Experimente. Reinbeck: Rowohlt.

Diekmann, Andreas. (2014). Die Anderen als sozialer Kontext. Zur Bedeutung strategischer Interaktion. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 66, 47-66.

Egas, Martijn, und Arno Riedl. (2008). The economics of altruistic punishment and the maintenance of cooperation. *Proceedings of Royal Society B,* 275, 871-878.

Elster, Jon. (2005). Fehr on Altruism, Emotion, and Norms. *Analyse und Kritik*, 27, S. 197-211. Stuttgart: Lucius und Lucius.

Elster, Jon. (2006). Altruistic Behavior and Altruistic Motivations. In: *Handbook of the Economics of Giving, Altruism and Reciprocity.* Hrsg. Serge-Christophe Kolm und Jean Mercier Ythier, Vol. 1, S. 183-206. Amsterdam: North-Holland (Elsevier),

Engel, Christoph. (2011). Dictator Games: A meta study. *Experimental Economics*, 14, 4:583-610.

Falk, Armin, Ernst Fehr, und Urs Fischbacher. (2003). On The Nature Of Fair Behavior. *Economic Inquiry*. 41, 1:20-26.

Fehr, Ernst, und Simon Gächter. (2000). Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives*, 14, 3:159-181.

Fehr, Ernst, und Simon Gächter. (2002). Altruistic punishment in humans. *Nature*, 415, 137-140

Fehr, Ernst, und Urs Fischbacher. (2002). Why Social Preferences Matter – The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives. *Economic Journal*, 112, 478:C1-C33.

Fehr, Ernst, und Urs Fischbacher. (2003). The nature of human altruism. *Nature,* 425, 785-791.

Fehr, Ernst, und Urs Fischbacher. (2004). Third-party punishment and social norms. *Evolution and Human Behavior*, 25, 63-87.

Fischbacher, Urs, Christina M. Fong, und Ernst Fehr. (2009). Fairness, Errors and the Power of Competition. *Journal of Economic Behavior und Organization*, 72, 527-545.

Fischbacher, Urs. (2007). z-Tree: Zurich Toolbox for Ready-made Economic Experiments. *Experimental Economics,* 10, 2:171-178.

Franzen, Axel, und Sonja Pointner. (2012). Anonymity in the dictator game revisited. *Journal of Economic Behavior und Organization*, 81, 74-81.

Gintis, Herbert, Samuel Bowles, Robert Boyd, und Ernst Fehr. (2003). Explaining altruistic behavior in humans. *Evolution and Human Behavior,* 24:153-172.

Grosskopf, Brit. (2003). Reinforcement and Directional Learning in the Ultimatum Game with Responder Competition. *Experimental Economics*, 6, 141–158.

Güth, Werner, Carsten Schmidt, und Matthias Sutter. (2007). Bargaining outside the lab – a newspaper experiment of a three-person ultimatum game. *The Economic Journal*, 117, 518:449-469.

Güth, Werner, Rolf Schmittberger, und Bern Schwarze. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior and Organization* 3, 4: 367-388.

Güth, Werner, und Eric van Damme. (1998). Information, Strategic Behavior, and Fairness in Ultimatum Bargaining: An Experimental Study. *Journal of Mathematical Psychology,* 42, 227-247.

Güth, Werner, und Martin Kocher. (2014). More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature. Jou*rnal of Economic Behavior and Organization*, 108, 396-409.

Halko, Marja-Liisa, und Timo Seppälä. (2006). Ultimatum Game Experiments. Discussion Paper No. 140.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin F. Camerer, Ernst Fehr, Herbert Gintis, und Richard McElreath. (2001). In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies. *The American Economic Review,* 91, 2:73-78.

Kagel, John H., und Katherine Willey Wolfe. (2001). Tests of Fairness Models Based on Equity Considerations in a Three-Person Ultimatum Game. *Experimental Economics*, 4, 203-219

Knez, Marc J., und Colin F. Camerer. (1995). Outside Options and Social Comparison in Three-Player Ultimatum Game Experiments. *Games and Economic Behavior,* 10, 65-94.

Loewenstein, George. (1999). Experimental Economics from the Vantage-point of Behavioural Economics. *The Economic Journal*, 109, 453:25–34.

Nikiforakis, Nikos. (2013). Self-Governance through altruistic punishment? In *Reward and Punishment in Social Dilemmas.* Hrsg. Paul Van Lange, Bettina Rockenbach, und Toshio Yamagishi. 197- 213. New York: Oxford University Press.

Nowak, Martin A., Karen M. Page, und Karl Sigmund. (2000). Fairness Versus Reason in the Ultimatum Game. *Science*, 289, 1773-1775.

Nowak, Martin A., und Karl Sigmund. (2005). Evolution of indirect reciprocity. *Nature*, 437, 1291-1298.

Ockenfels, Axel. (1999). Fairness, Reziprozität und Eigennutz. Ökonomische Theorie und experimentelle Evidenz. Tübingen: Mohr Siebeck.

Okade, Akira, und Arno Riedl. (2005). Inefficiency and social exclusion in a coalition formation game: experimental evidence. *Games and Economic Behavior*, 50, 278-311.

Oppewal, Harmen, und Elena Tougareva. (1992). A three-person ultimatum game to investigate effects of differences in need, sharing rules and observability on bargaining behaviour. *Journal of Economic Psychology,* 13, 203-213.

Pointner, Sonja. (2012). Das Fairnesskalkül. Robustheit, Determinanten und externe Validität der Fairnessnorm. Wiesbaden: VS Springer.

Przepiorka, Wojtek, und Andreas Diekmann. (2013). Individual heterogeneity and costly punishment: a volunteer's dilemma. *Proceedings of The Royal Society B,* 280.

Riedl, Arno, und Jana Výrašteková. (2002). Social Preferences in Three-Player Ultimatum Game Experiments. Tilburg University. Center for Economic Research. Discussion paper.

Rushton, J Philippe, Roland D. Chrisjohn, und Cynthia Fekken. (1981). The altruistic personality and the self-report altruism scale. *Personality and Individual Differences*, 2, 293-302.

Schmitt, Pamela, Robert Schupp, Kurtis Swope, und Justian Mayer. (2008). Pre-commitment and personality: Behavioral explanations in ultimatum games. *Journal of Economic Behavior and Organization*, 66, 3-4:597–605.

Vieth, Manuela. (2003). Die Evolution von Fairnssnormen im Ultimatumspiel. Eine spieltheoretische Modellierung. *Zeitschrift für Soziologie*, 32, 4:346-367.

Voss, Thomas. (2001). Game Theoretical Perspectives on the Emergence of Social Norms. In *Social Norms*. Hrsg. Michael Hechter, und Karl-Dieter Opp. 105-136. New York: Russell Sage.

Weber, J. Mark, Shirli Kopelman, und David M. Messick. (2004). A Conceptual Review of Decision Making in Social Dilemmas: Applying a Logic of Appropriateness. *Personal and Social Psychology Review*, 8, 3:281-307.

Zimbardo, Philip, und Richard Gerrig. (2008). Psychologie. München: Pearson Studium, 18. aktualisierte Auflage.

Anhang

1) Deskriptive Übersicht unabhängiger Variablen der logistischen Regression (LEx), n=115

Variablen	Mittelwert	Standardabweichung	Min	Max
Haupteffekte				
eigene Abgaben	2.16	0.71	0.1	5
Fingiertes Angebot	0.50	0.50	0	1
Design (4R)	0.56	0.50	0	1
Showup-fee	0.75	0.44	0	1
Kontrollvariablen				
Alter	23.7	4.02	19	40
Anonymitätsgefühl	5.38	1.14	0	6
Teilnahme	0.24	0.43	0	1
Altruismus	28.4	7.08	9	42

2) Deskriptive Übersicht unabhängiger Variablen der logistischen Regression (Bern)

Variablen	n	Mittelwert	Standardabweichung	Min	Max
Haupteffekte					
eigene Abgaben (multipler Responder)	81	3.84	1.05	1	5
eigene Abgaben (Kontrollgruppe)	36	4.01	1.50	0	5
eigene Abgaben (insgesamt)	117	3.89	1.21	0	5
Multiple Responder	117	0.69	0.46	0	1
Kontrollvariablen					
Weibliches Geschlecht	114	0.64	0.48	0	1
Fingiertes Angebot	117	0.45	0.49	0	1
Wirtschaftsstudium	117	0.24	0.43	0	1
Alter	117	23.7	2.70	19	34
Einkommen	117	617.48	683.73	0	3800
Anonymitätsgefühl	117	6.36	1.06	2	7
Teilnahme	117	0.23	0.42	0	1

3) Spieleranweisung

Folgend werden die ursprünglichen Formulierung sowie die geänderte Fassung der Spieleranweisungen des simultanen Designs inklusive der Quizfrage in der Reproduktion aufgeführt. Weiter unten findet sich der Vollständigkeit halber die Anleitung für die Situation sequentieller Responder.

Ursprüngliche Formulierung der ersten Spielsituation:

„**Liebe Teilnehmerin, lieber Teilnehmer**

eine Person an einem der anderen PCs wird Ihnen nun einen Vorschlag machen, wie ein Betrag von 10 Franken zwischen ihr/ihm und Ihnen aufgeteilt werden soll.

Sie können den Vorschlag akzeptieren oder ablehnen. Falls Sie die Aufteilung akzeptieren, dann erhält jeder den für ihn vorgesehenen Betrag. Falls Sie ablehnen, dann erhalten Sie, aber auch der Mitspieler, keine Auszahlung.

Der Vorschlag des ersten Spielers geht gleichzeitig auch noch an einen zweiten weiteren Mitspieler hier in diesem Raum. Dieser Mitspieler befindet sich in der gleichen Entscheidungssituation wie Sie. Er kann den Betrag akzeptieren oder ablehnen. Akzeptiert er, würde er den Auszahlungsbetrag erhalten. Lehnt er ab, erhält er keinen Auszahlungsbetrag.

Sollten Sie beide akzeptieren, dann erhält jeder von ihnen den vom ersten Spieler vorgeschlagenen Auszahlungsbetrag. Gleichzeitig erhält auch der erste Spieler den Betrag, den er für sich selbst vorgesehen hat. Lehnen beide Empfänger ab (also Sie und der zweite Mitspieler), dann erhalten weder die Empfänger noch der erste Spieler den Betrag. Wenn einer der Empfänger den Auszahlungsbetrag ablehnt, erhalten der erste Spieler und der Empfänger, der abgelehnt hat, nichts.

Sie werden nicht erfahren, von welcher Person dieser Vorschlag kommt. In Kürze werden Sie den Vorschlag für die Aufteilung der 10 Franken sehen.“

Geänderte Fassung mit zwei (*vier*) Respondern:

„**Liebe Teilnehmerin, lieber Teilnehmer**

eine Person (P) an einem der anderen PCs wird Ihnen nun einen Vorschlag machen, wie ein Betrag von 5 Euro zwischen ihr/ihm (P) und Ihnen (R-Spieler) aufgeteilt werden soll.

Sie können den Vorschlag akzeptieren oder ablehnen. Falls Sie die Aufteilung akzeptieren, dann erhalten Sie und auch P den vorgesehenen Betrag. Falls Sie ablehnen, dann erhalten weder Sie noch P etwas.

Der Vorschlag von P geht zur selben Zeit an einen *(drei)* weiteren Spieler. Dieser befindet sich in der gleichen Entscheidungssituation wie Sie.
Sobald ein R-Spieler ablehnt, erhält dieser und auch P nichts. Jeder R-Spieler, der das Angebot

annimmt, bekommt seinen Anteil.

Sie werden nicht erfahren, von welcher Person dieser Vorschlag kommt. In Kürze werden Sie den Vorschlag für die Aufteilung der 5 Euro sehen.“

Ursprüngliche Präsentation Angebot und Entscheidung:

„Spieler 1 macht Ihnen folgenden Vorschlag:

Spieler 1 möchte 9 Franken behalten

Sie bekommen 1 Franken

Entscheiden Sie nun, ob Sie den Vorschlag akzeptieren: …“

Geänderte Fassung Präsentation Angebot und Entscheidung:

„Der Spieler P macht Ihnen folgenden Vorschlag:

Spieler P möchte 4 Euro behalten

Sie bekommen 1.00 Euro.

Entscheiden Sie nun, ob Sie den Vorschlag akzeptieren:…“

Zweite Spielsituation:

Ursprüngliche Offerte des Probanden:

„Jetzt können **Sie** einen Betrag von 10 Franken aufteilen.

Zwei Spieler in diesem Raum erhalten Ihren Aufteilungsvorschlag. Wenn mindestens einer der Mitspieler akzeptiert, bekommen Sie den Betrag, den Sie für sich selbst vorgesehen haben. Wenn keiner Ihrer Mitspieler akzeptiert, bekommen Sie keinen Betrag aus dieser Entscheidungssituation.

Jeder Ihrer Mitspieler, der akzeptiert, bekommt den ihm zUltimatumspielewiesenen Betrag. Falls beide akzeptieren, bekommen auch beide Mitspieler den Betrag.

Sie können jede beliebige Aufteilung der 10 Franken vornehmen. Bitte tragen Sie Ihren Aufteilungsvorschlag in die vorgegebenen Felder ein.

Tragen Sie in dieses Kästchen den Betrag ein, den Sie gerne behalten möchten:
Tragen Sie in dieses Kästchen den Betrag ein, den Sie dem Mitspieler zuteilen:“

Geänderte Fassung mit zwei (*vier*) Respondern:

„Jetzt können Sie einen Betrag von 5 Euro aufteilen (P-Spieler).

Zwei *(vier)* andere Spieler (R-Spieler) in diesem Raum erhalten Ihren Aufteilungsvorschlag. Sobald einer davon das Angebot ablehnt, bekommen Sie nichts. Sobald einer davon akzeptiert, bekommen Sie den Betrag, den Sie für sich selbst vorgesehen haben.

Jeder der R-Spieler, der akzeptiert, bekommt den ihm zugewiesenen Betrag. Falls beide akzeptieren, bekommen auch beide Mitspieler den Betrag. Sie können jede beliebige Aufteilung der 5 Euro vornehmen. Bitte runden Sie auf 10 Cent Genauigkeit

Tragen Sie in dieses Kästchen den Betrag ein, den Sie gerne behalten möchten:
Tragen Sie in dieses Kästchen den Betrag ein, den jeder R-Spieler erhalten soll:“

Obligatorische Quizfrage:

„Bitte beantworten Sie folgende Frage:
Ein Ball und ein Schläger kosten zusammen 1.10 Euro. Der Schläger kostet 1.00 Euro mehr als der Ball. Wie viel kostet der Ball?“

Spielanleitung zum Design mit sequentiellen Respondern („Konkurrenzsituation“):

Für den Responder 1:

„Lieber Teilnehmerin, lieber Teilnehmer,

eine Person an einem der anderen PCs wird Ihnen nun einen Vorschlag machen, wie ein Betrag von 10 Franken zwischen ihr/ihm aufgeteilt werden soll.
Sie werden nicht erfahren, von welcher Person dieser Vorschlag kommt. Sie werden gleich den Vorschlag für die Aufteilung der 10 Franken sehen.“

„„Spieler 1 macht Ihnen folgenden Vorschlag:

Spieler 1 möchte 8 Franken behalten

Sie bekommen 2 Franken.

Sie können sich nun überlegen, ob Sie die Entscheidung der Person akzeptieren oder ablehnen.

Falls Sie mit der Aufteilung einverstanden sind, dann akzeptieren Sie die Entscheidung. Sie werden dann den Betrag erhalten, der Ihnen zugeteilt worden ist. Falls Sie mit der Aufteilung nicht einverstanden sind, haben Sie die Möglichkeit, ein Veto einzulegen. In diesem Fall wird dem ersten Spieler mitgeteilt, dass Sie die Entscheidung ablehnen. Sie und der vorschlagende Mitspieler erhalten dann keinen Auszahlungsbetrag.

Der erste Spieler hat dann die Möglichkeit, den Aufteilungsvorschlag an einen anderen Spieler hier im Raum weiterzuleiten. Auch dieser Spieler hat, wie Sie, die Möglichkeit, den Vorschlag zu

akzeptieren oder abzulehnen. Akzeptiert der zweite Spieler die Aufteilung von Spieler 1, dann erhalten die beiden Spieler den für sie vorgesehenen Betrag. Falls der zweite Spieler den Vorschlag ablehnt, dann erhält keiner der Spieler eine Auszahlung.

Alle erspielten Beträge werden später in realen Beträgen ausgezahlt.

Entscheiden Sie nun, ob Sie den Vorschlag akzeptieren: …“

Für den Responder 2:

„Ein Spieler aus diesem Raum hat in der vorherigen Runde einem Ihrer Mitspieler ein Angebot macht, das von diesem abgelehnt wurde. Der erste Spieler möchte dieses Angebot nun Ihnen machen.

Der Mitspieler möchte 9 Franken behalten

Sie bekommen 1 Franken

Sie können sich nun überlegen, ob Sie die Entscheidung der Person akzeptieren oder ablehnen.

Falls Sie mit der Aufteilung einverstanden sind, dann akzeptieren Sie die Entscheidung. Sie werden dann den Betrag erhalten, der Ihnen zugeteilt worden ist. Spieler 1 wird ebenfalls den Betrag erhalten, den er für sich vorgesehen hat. Der Mitspieler, der das Angebot von Spieler 1 in der ersten Runde abgelehtn hat, wird dagegen nichts bekommen.

Falls Sie mit der Aufteilung nicht einverstanden sind, lehnen Sie das Angebot ab. In diesem Falle werden weder Sie noch Spieler 1 eine Auszahlung erhalten.

Alle erspielten Beträge werden später in realen Beträgen ausgezahlt.

Entscheiden Sie nun, ob Sie den Vorschlag akzeptieren: …“

Als Proposer:

„Jetzt können **Sie** einen Betrag von 10 Franken aufteilen.

Ein Mitspieler aus diesem Raum wird Ihren Aufteilungsvorschlag erhalten. Wenn der Mitspieler akzeptiert, bekommen Sie den Betrag, den Sie für sich selbst vorgesehen haben und Ihr Mitspieler den ihm zugewiesenen Betrag. Wenn Ihr Mitspieler ablehnt, dann bekommt ein weiterer Spieler aus diesem Raum Ihren Aufteilungsvorschlag. Wenn dieser Spieler auch ablehnt, bekommen Sie keinen Betrag aus dieser Entscheidungssituation. Sie können jede beliebige Aufteilung der 10 Franken vornehmen. Bitte tragen Sie ihren Aufteilungsvorschlag in die vorgegebenen Felder ein.

Tragen Sie in dieses Kästchen den Betrag ein, den Sie gerne behalten möchten:

Tragen Sie in dieses Kästchen den Betrag ein, den Sie dem/der Mispieler/in zuteilen:…“

4) Fragebogen im Anschluss an das Experiment in Bern

Rekonstruktion anhand des Codebuchs. Das Original wurde mittels „Unipark“ programmiert und durchgeführt.

Haben Sie schon einmal an einem ähnlichen Experiment teilgenommen?
1 nein, ich habe noch nie an einem solchen Experiment teilgenommen
2 ja, ich habe schon an einem ähnlichen Experiment teilgenommen
3 ja, ich habe schon an mehreren ähnlichen Experimenten teilgenommen

Denken Sie bitte noch einmal an die Entscheidungssituation im Experiment zurück! Wie hoch war das erste Angebot, dass (*sic!*) Sie von Ihrem Mitspieler erhalten haben?

Und wie hätte das Angebot aussehen müssen, damit Sie es gerade noch als akzeptabel empfinden?

Bitte geben Sie an, wie anonym Ihnen die Entscheidungssituation vorgekommen ist.
1 überhaupt nicht anonym
2
3
4
5
6
7 vollständig anonym

Wie verständlich fanden Sie die Instruktionen in unserem Experiment?

1 Ich habe nichts verstanden.
2
3
4
5
6
7 Ich habe alles verstanden.

Jetzt geht es um allgemeine Fragen zum Thema Gesellschaft. Man kann unterschiedliches Vertrauen in gesellschaftliche Einrichtungen haben. Bitte geben Sie für jede der folgenden Einrichtungen an, wie viel Vertrauen Sie in diese haben!

Vertrauen in...	1 sehr viel Vertrauen	2 eher viel Vertrauen	3 weder besonders viel noch besonders wenig Vertrauen	4 eher wenig Vertrauen	5 sehr wenig Vertrauen
den Bundesrat					
die politischen Parteien					
das Rechtssystem					
Wirtschaftsunternehmen					
die öffentliche Verwaltung					
Medien					
Kirchen					
die Gewerkschaft					
das Bildungssystem					
die Polizei					

Auch gegenüber den Mitmenschen kann man unterschiedlich vertrauensvoll sein. Wie würden Sie ihre Einstellung zu anderen Personen beschreiben?

	1 stimme voll zu	2 stimme eher zu	3 weder zutreffend noch unzutreffend	4 lehne eher ab	5 lehne voll ab
Im Allgemeinen kann man den Menschen vertrauen.					
Heutzutage kann man sich auf niemanden mehr verlassen.					
Wenn man mit Fremden zu tun hat, ist es besser, vorsichtig zu sein, bevor man ihnen vertraut.					

Bitte geben Sie an, inwieweit die Aussagen auf Sie zutreffen:

	1 trifft völlig zu	2 trifft eher zu	3 trifft eher nicht zu	4 trifft gar nicht zu
Mein erster Eindruck von Menschen stellt sich gewöhnlich als richtig heraus.				
Ich weiß immer genau, wieso ich etwas mag.				
Viele Leute halten mich für außergewöhnlich.				
Ich ziehe manchmal den Kürzeren, weil ich mich nicht schnell genug entscheiden kann.				
Ich bin mir sehr sicher in meinem Urteil.				
Zur Not lüge ich auch schon mal.				

Und was sagen Sie zu den weiteren Aussagen?

	1 trifft völlig zu	2 trifft eher zu	3 trifft eher nicht zu	4 trifft gar nicht zu
Ich habe noch niemals etwas im Geschäft oder anderswo beschädigt, ohne es zu melden.				
Ich sage immer, was ich denke.				
Ich bin manchmal ärgerlich, wenn ich meinen Willen nicht bekomme.				
Ich bin immer gewillt, einen Fehler, den ich mache, auch zuzugeben.				
Ich habe gelegentlich mit Absicht etwas gesagt, was die Gefühle des anderen verletzen könnte.				

Im nächsten Fragenblock geht es um das Thema persönliche Netzwerke. Wir möchten Sie bitten, an Ihre sozialen Kontakte, also an Ihre Familie, Ihre Verwandten und Ihre Freunde, zu denken. Nennen Sie im Folgenden bitte die Vornamen (oder andere Abkürzungen) der Personen, deren Rat Sie einholen, wenn Sie wichtige Entscheidungen in Ihrem Leben Treffen:
Person 1
Person 2
Person 3
Person 4
Person 6

Bitte geben Sie das Geschlecht Ihres Netzwerkkontaktes an:

1 weiblich
2 männlich

Wie alt ist Ihr Netzwerkkontakt?

In welcher Beziehung stehen Sie zu der Person:
1 Ehemann/Ehefrau bzw. Lebensgefährte/in
2 Vater
3 Mutter
4 Schwester/Bruder (auch Halbgeschwister)
5 Tante/Onkel
6 Großmutter/Großvater
7 andere/r Verwandte/r
8 beste/r Freund/in
9 Freund/in bzw. Kollege/in
10 Arbeitskollege/in
11 Nachbar/in
12 entfernter Bekannte/r
13 Sonstiges:
Sonstiges:

Wie eng ist die Beziehung zu diesem Kontakt?
1 sehr eng
2 ziemlich eng
3 weder noch
4 nicht besonders eng
5 gar nicht eng

Welchen Beruf übt Ihr Netzwerkkontakt momentan aus?

Bitte nennen Sie den höchsten schulischen Bildungsabschluss Ihres Netzwerkkontaktes:
1 Kein Abschluss
2 Obligatorische Schule
3 Anlehre
4 Diplommittelschule, Handelsschule, Haushaltslehrjahr
5 Berufslehre, BMS
6 Matura, Lehrerseminar
7 Meisterdiplom, Eidgenössischer Fachausweis, Techniker oder Fachschule
8 Universität
9 noch in Ausbildung
10 weiß nicht
11 Höhere Fachschule / Fachhochschule (HTL, FHS,HWV etc.)

Wie viele Personen würden Sie zu Ihrem engen Freundeskreis zählen (Verwandte nicht mitgezählt)?

Wie viele Personen würden Sie zu Ihrem erweiterten Freundeskreis zählen (gemeint sind Personen, die Sie persönlich kennen, zu denen aber nur ein schwacher Kontakt besteht) (Verwandte nicht mitgezählt)?

Zu wie vielen Ihrer Verwandten (inklusive Ihrer Familienmitglieder) pflegen Sie enge Kontakte?

Und zu wie vielen Ihrer Verwandten haben Sie eher nur schwachen Kontakt?

Was würden Sie sagen, wenn Sie an das Verhältnis zu Ihren Eltern denken: Wie eng ist die Beziehung zu...

Ihrer Mutter?
1 sehr eng
2 ziemlich eng
3 weder noch
4 nicht besonders eng
5 gar nicht eng
6 keine Angabe möglich

Ihrem Vater?
1 sehr eng
2 ziemlich eng
3 weder noch
4 nicht besonders eng
5 gar nicht eng
6 keine Angabe möglich

Abschließend haben wir noch ein paar Fragen zur Sozialstatistik. Haben Sie Geschwister?
1 ja
2 nein

Wann sind diese geboren?
Geschwister 1
Geschwister 2
Geschwister 3
Geschwister 4
Geschwister 5

Wenn Sie nun an das Verhältnis zu Ihren Geschwistern denken: Wie eng ist das Verhältnis zu dem in dem folgenden Jahr geborenen Geschwister?

	1 sehr eng	2 ziemlich eng	3 weder noch	4 nicht besonders eng	5 gar nicht eng
Geschwister 1					
Geschwister 2					
Geschwister 3					
Geschwister 4					
Geschwister 5					

Sie sind...
1 weiblich
2 männlich

In welchem Jahr sind Sie geboren?

Bitte geben Sie an, wie hoch Ihr monatliches Einkommen ist, das für Sie frei verfügbar ist (abzüglich von Miete, Lebenshaltungskosten, etc.). Falls Sie Ihr verfügbares Einkommen nicht genau angeben können oder es monatlich schwankt, dann tragen Sie bitte eine Schätzung ein.

Welches Hauptfach studieren Sie?

Welche Nebenfächer studieren Sie?

Wenn Sie auch in Zukunft Einladungen zu Experimenten bekommen möchten, dann geben Sie bitte hier E-Mail-Adresse an:

5) Fragebogen im Anschluss an das Experiment in Leipzig

Ihnen werden jetzt ein paar Fragen zu Ihrer Person gestellt. Bitte antworten Sie wahrheitsgemäß, Ihre Angaben bleiben anonym.

Wie alt sind Sie (in Jahren)?

Welches Geschlecht haben Sie?
0 = weiblich
1 = männlich

Studieren Sie im Moment?
1 = ja
0 = nein

Haben Sie Geschwister?
0=nein
1=ja

Zu welchem Abschluss führt der Studiengang, in dem Sie im Moment eingeschrieben sind?
1 Bachelor
2=Master
3=Diplom
4=Magister
5=Staatsexamen
6=Promotion
7=Anderer

Wieviele Semester sind Sie in Ihrem aktuellen Studiengang schon eingeschrieben?

Was ist das Studienfach in Ihrem aktuellen Studiengang, mit dem Sie sich bisher am meisten beschäftigt haben?
1=Soziologie
2=Politikwissenschaft
3=Mathematik
4=Physik
5=BWL/VWL
6=Informatik
7=Ingenieurwissenschaft
8=Psychologie
9=Sprachwissenschaft
10=Geschichte
99=Nichts davon

Haben Sie schon jemals ein Buch gelesen, das sich überwiegend mit Spieltheorie oder Logik beschäftigt?
1=ja
0=nein

Haben Sie schon jemals einen Kurs (Vorlesung, Seminar, ...) besucht, der sich überwiegend mit Spieltheorie oder Logik beschäftigt hat?
1=ja
0=nein

Haben Sie, abgesehen von der heutigen Sitzung, schon an Laborexperimenten teilgenommen?
1=ja
0=nein

Bitte geben Sie an, wie anonym Ihnen die Entscheidungssituation vorgekommen ist:
1 überhaupt nicht anonym
2
3
4
5
6
7 vollständig anonym

Wie verständlich fanden Sie die Instruktionen in unserem Experiment?
1 Ich habe nichts verstanden.
2
3
4
5
6
7 Ich habe alles verstanden.

Denken Sie bitte noch einmal an die Entscheidungssituation im Experiment zurück. Wie hoch war das erste Angebot, das Sie von Ihrem Mitspieler erhalten haben?

Wie viele andere R-Spieler gab es neben Ihnen?

Wie hätte das Angebot des P-Spielers aussehen müssen, damit Sie es gerade noch als akzeptabel empfinden?

In Ihrer Rolle als P-Spieler: Was glaubten Sie, wie viele der R-Spieler Ihr Angebot ablehnen würden? (*Angabe über einen Schieberegler***)**
1 niemand
2
3
4
5
6
7
8
9
10 alle

In Ihrer Rolle als R-Spieler: Was glaubten Sie, wie viele der anderen R-Spieler das Angebot ablehnen würden? (Angabe über einen Schieberegler)
1 niemand
2
3
4
5
6
7
8
9
10 alle

Bitte geben Sie Ihre Zustimmung zu folgenden Aussagen ab:

	1 stimme voll zu	2 stimme eher zu	3 weder zutreffend noch unzutreffend	4 lehne eher ab	5 lehne voll ab
Im Allgemeinen kann man den Menschen vertrauen.					
Heutzutage kann man sich auf niemanden mehr verlassen.					
Wenn man mit Fremden zu tun hat, ist es besser, vorsichtig zu sein, bevor man ihnen vertraut.					

Im Anschluss an den Fragebogen bis dato folgten einige Items der Self-Report-Altruism-Scale (Rushton et al.. 1981). Ausgewählte und im Fragebogen implementierte Items hervorgehoben durch die deutsche Übersetzung in Klammern.

Kreuzen Sie bitte die zutreffende Kategorie an, wie oft Sie folgendes bereits getan haben:

	Never (*niemals*)	Once (*einmal*)	More than once (*mehr als einmal*)	Often (*häufig*)	Very often (*sehr häufig*)
I have helped push a stranger's car out of the snow.					
I have given directions to a stranger. (*Ich habe einem Fremden den Weg beschrieben*)					
I have made change for a stranger. (*Ich habe einem Fremden Geld gewechselt*)					
I have given money to a charity. (*Ich habe Geld zu wohltätigen Zwecken gespendet*)					
I have given money to a stranger who needed it (or asked me for it). (*Ich habe einem Fremden Geld gegeben, der es benötigte bzw. danach fragte*)					
I have donated goods or clothes to a charity. (*Ich habe Kleider und andere Artikel zu wohltätigen Zwecken gespendet*)					
I have done volunteer work for a charity. (*Ich habe für eine wohltätige Organisation gearbeitet*)					
I have donated blood. (*Ich habe Blut gespendet*)					
I have helped carry a stranger's belongings (books, parcels, etc.). *(Ich habe Fremden geholfen, Pakete, Bücher oder Ähnliches zu tragen)*					
I have delayed an elevator and held the door open for a stranger. (*Ich habe einen Aufzug verzögert bzw. eine Tür für Fremde offen gehalten*)					
I have allowed someone to go ahead of me in a lineup (at photocopy machine, in the supermarket). (*Ich habe jemanden in der Warteschlange vorgelassen*)					

I have given a stranger a lift in my car.					
I have pointed out a clerk's error (in a bank, at the supermarket) in undercharging me for an item. (*Ich habe eine Kassiererin darauf hingewiesen, dass ich zu viel Rückgeld bekommen habe*)					
I have let a neighbour whom I didn't know too well borrow an item of some value to me (e.g., a dish, tools, etc.).					
I have bought 'charity' Christmas cards deliberately because I knew it was a good cause.					
1 have helped a classmate who 1 did not know that well with a homework assignment when my knowledge was greater than his or hers.					
I have before being asked, voluntarily looked after a neighbour's pets or children without being paid for it.					
I have offered to help a handicapped or elderly stranger across a street. (*Ich habe behinderten oder älteren Fremden meine Hilfe angeboten)*					
I have offered my seat on a bus or train to a stranger who was standing. (*Ich habe Fremden in Bus oder Bahn meinen Sitz angeboten*)					
I have helped an aquaintance to move households.					