

Prof. Dr. Cornelius Prittwitz/Prof. Dr. Michael Baurmann/  
Prof. Dr. Klaus Günther/Prof. Dr. Matthias Jahn/  
Prof. Dr. Lothar Kuhlen/Prof. Dr. Reinhard Merkel/  
Prof. Dr. Cornelius Nestler/Prof. Dr. Lorenz Schulz (Hrsg.)

## Rationalität und Empathie

Kriminalwissenschaftliches Symposium für Klaus Lüderssen  
zum 80. Geburtstag



Nomos

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

ISBN 978-3-8487-1252-6

1. Auflage 2014

© Nomos Verlagsgesellschaft, Baden-Baden 2014. Printed in Germany. Alle Rechte,  
auch die des Nachdrucks von Auszügen, der fotomechanischen Wiedergabe und der  
Übersetzung, vorbehalten. Gedruckt auf alterungsbeständigem Papier.

## Strafen aus Spaß?

### Experimentelle Befunde zum Sanktionsverhalten

*Michael Baumann*

Ich diskutiere im Folgenden drei Varianten von Experimenten, die sich mit menschlichem Sanktionsverhalten beschäftigen. „Sanktionsverhalten“ ist hier in einem weiten Sinn verstanden, nicht nur als negative Reaktion, die aufgrund eines unerwünschten Verhaltens erfolgt, sondern auch als Übelzufügung, die nicht durch ein unterstelltes Fehlverhalten verursacht wurde. Zwischen den Experimenten besteht insofern nur eine Familienähnlichkeit, sie thematisieren unterschiedliche Kontexte und Verhaltensweisen. Dennoch denke ich, dass sich einige verallgemeinernde Schlussfolgerungen ziehen lassen, die auch für das Gebiet staatlichen Strafens von Bedeutung sind.

Die Experimente unterscheiden sich ebenfalls erheblich im Grad ihrer Bekanntheit: Eine der Experimentvarianten ist über ein Fachpublikum hinaus auch in der Öffentlichkeit zum Thema geworden, über eine andere existiert zumindest innerhalb der Fachwelt ein verbreitetes Wissen. Die dritte Variante wird schließlich unbekannt sein. Mit ihr will ich beginnen. Ich stelle zunächst die Experimente und ihre Ergebnisse dar, anschließend diskutiere ich unterschiedliche Möglichkeiten der Interpretation und Erklärung dieser Ergebnisse.

#### *1. Ergebnisse der Experimente*

##### *1.1 Joy of Destruction-Experimente*

Bei den Joy of Destruction-Experimenten handelt es sich um Experimente, die von einer Gruppe von deutschen Forschern in der jüngeren Vergangenheit in Namibia durchgeführt wurden. Die Experimente und ihre Ergebnisse wurden im November 2011 auf einem Kolloquium am ZiF in Bielefeld zu

Ehren von Elinor Ostrom vorgestellt.<sup>1</sup> Mit den Experimenten sollte der Einfluss von Ressourcenknappheit auf Sozialverhalten untersucht werden. In diesem Zusammenhang ist die Hypothese populär – nicht zuletzt aufgrund der Studien von Ostrom (1990) –, dass es eine positive Korrelation zwischen Knappheit und prosozialem Verhalten gibt: Da Knappheit und Mangel durch Kooperation und Zusammenarbeit prima facie besser bewältigt werden können als durch isolierte Anstrengungen und Konfliktstrategien, scheint die Annahme plausibel, dass in Kontexten mit einer suboptimalen Ressourcenausstattung eher mehr gemeinsames Handeln zu finden ist als in Kontexten, in denen Wohlstand herrscht. Kurz gesagt lautet die Hypothese: Armut fördert Kooperation.

In den Joy of Destruction-Experimenten wird die entgegengesetzte Hypothese geprüft. Kann Ressourcenknappheit möglicherweise auch zu einem antisozialen Verhalten führen? Können Mangel und Not zu einem Bedürfnis oder sogar Vergnügen führen, anderen zu schaden, ohne dafür materielle persönliche Vorteile zu erzielen oder Vergeltung für ein Fehlverhalten zu üben? Zugespitzt lautet die Gegenhypothese: Armut fördert Boshaftigkeit.

Das Design der Experimente, um diese unerfreuliche Hypothese zu prüfen, ist denkbar einfach. In einer Region mit halbtrockenem Weideland in Namibia wurden zwei Gebiete lokalisiert, die sich signifikant in ihrer Produktion an Biomasse unterscheiden, ein Gebiet mit relativ hoher Produktivität und ein Gebiet mit niedriger Produktivität. Der Unterschied ist erklärbar mit der jeweiligen Bodenqualität. Die Verfügbarkeit über lebenswichtige Ressourcen in Form vor allem des Viehbestandes unterscheidet sich entsprechend in den beiden Gebieten. Der kulturelle und politische Hintergrund ist jedoch für die Bewohner der beiden Gebiete identisch.

In jedem Gebiet wurden für die Experimente 10 Sitzungen mit insgesamt 60 Versuchspersonen durchgeführt. Jeweils zwei Personen wurden nach dem Zufallsprinzip für ein Experiment ausgewählt. Sie blieben anonym und hatten auch keine Gelegenheit, ihre Partner in dem Experiment im Nachhinein zu identifizieren. Das Experiment bestand jeweils nur aus einer Runde. Strategische Überlegungen, Lern- und Reputationseffekte waren damit ausgeschlossen.

Beide Teilnehmer erhielten jeweils N\$ 10,- als anfängliche Ausstattung. Sie konnten sich entscheiden, ob sie den Betrag unangetastet lassen oder N

---

<http://www.uni-bielefeld.de/ZiF/AG/2011/11-17-Dawid.pdf>. Prediger/Vollan/Herrmann (2013).

\$ 1,- dafür ausgeben, damit die N\$ 10,- ihres Partners um N\$ 5,- reduziert werden. Eine solche „Zerstörung“ erbrachte demnach keine materiellen Vorteile für den „Zerstörer“, sondern war im Gegenteil mit monetären Kosten verbunden.

Drei Ergebnisse waren möglich: Erstens konnten beide Teilnehmer entscheiden, das Einkommen des anderen nicht zu reduzieren. Beide würden dann N\$ 10,- behalten. Dieses Ergebnis würde man erwarten, wenn beide Teilnehmer ausschließlich ihren materiellen Gewinn maximieren wollen. Zweitens konnten beide entscheiden, den Geldbetrag des anderen zu reduzieren, in diesem Fall würden beide mit N\$ 4,- nach Hause gehen. Drittens konnte einer der beiden entscheiden, den Geldbetrag des anderen zu reduzieren, während der andere auf eine „Zerstörung“ verzichtet; in diesem Fall würde der „Zerstörer“ N\$ 9,- erhalten, während das „Opfer“ lediglich N\$ 5,- behielte.

Die tatsächlichen Ergebnisse der Experimente sind eindeutig: In der relativ wohlhabenden Gegend entschieden sich 23% der Versuchspersonen dafür, den Geldbetrag ihrer Partner zu kappen. In der ärmeren Gegend stieg der Anteil der „Zerstörer“ dagegen auf 40%. Die Korrelation dieser Verhaltensunterschiede mit der Ressourcenverfügbarkeit erweist sich als robust auch bei der Kontrolle anderer Variablen wie sozio-demographische Merkmale oder Gruppenstrukturen. Die Experimente zeigen demnach, dass Personen aus der Gegend mit knapperen Ressourcen in signifikant höherem Ausmaß dazu tendieren, andere Bewohner ihrer Gegend zu schädigen, als Personen, die in der Gegend mit geringerer Ressourcenknappheit leben.

Wie bereits angemerkt, geht es bei der Fragestellung, die diesen Joy of Destruction-Experimenten zugrunde liegt, nicht um Sanktionsverhalten im engeren Sinn. Es geht nicht um Strafe als Vergeltung oder Prävention, sondern um eine augenscheinlich bedingungslose Bereitschaft, anderen Personen unter Inkaufnahme eigener Kosten zu schaden und ihnen ein Übel zuzufügen. Die geschädigten Personen hatten gegenüber dem anderen Versuchsteilnehmer weder ein unerwünschtes Verhalten gezeigt, noch konnten sie von einem solchen Verhalten durch die Übelszufügung in Zukunft abgehalten werden. Das Verhalten in solchen Experimenten wird scheinbar folgerichtig auch nicht als strategisches und rationales Verhalten beschrieben und interpretiert, sondern als Ausdruck von Emotionen wie Zerstörungslust, Neid, Boshafigkeit, Missgunst oder Schadenfreude.

Doch auch wenn es sich bei den in diesen Experimenten dokumentierten Verhaltensweisen nicht um Sanktionen und Strafen im engeren Sinn handelt, so ist die Frage nach den verursachenden Dispositionen für das beobachtete

Verhalten auch für ein Verständnis von Strafen und Sanktionen relevant. Sollte es der Fall sein, dass es möglicherweise nicht nur ein emotional verankertes Straf- und Vergeltungsbedürfnis gibt, das durch unerwünschte Verhaltensweisen aktualisiert wird, sondern dass eine latente Lust an der Schädigung anderer Personen existiert, die auch dann handlungsmotivierend sein kann, wenn kein besonderer Anlass vorliegt, hat das auch für Strafen und Sanktionen beachtenswerte Konsequenzen.

Zum einen müsste man dann nicht nur emotionalisierte Vergeltungsbedürfnisse einkalkulieren, sondern möglicherweise auch quasi archaische Emotionen wie Schädigungslust und Freude an Übelszufügung. Zum anderen würde das „rationalistische Vorurteil“ relativiert, dass ein zentrales Problem der zuverlässigen Verhängung von Strafen und Sanktionen in ihrer Kostspieligkeit für die Strafenden und Sanktionierenden besteht. Wenn es stimmt, dass Menschen sogar Kosten in Kauf nehmen, um andere zu schädigen, die sich nichts haben zuschulden kommen lassen, dann braucht man sich um dieses „Problem“ kaum den Kopf zu zerbrechen.

Die zweite Schlussfolgerung legen auch die beiden folgenden Experimente nahe, wenn auch aus diametral unterschiedlichen Perspektiven.

## 1.2 Milgram-Experimente

Die Milgram-Experimente gehören zu den bekanntesten Verhaltensexperimenten überhaupt. Ihre Popularität hängt vor allem damit zusammen, dass sie in einem noch größeren Ausmaß irritierend sind als die eben vorgestellten Joy of Destruction-Experimente.

Seit den ersten Experimenten, die Stanley Milgram in den 1960er Jahren durchführte, ist sein experimentelles Design in unzähligen Varianten in allen möglichen gesellschaftlichen und kulturellen Kontexten repliziert worden. Das Fazit aus allen diesen Experimenten konvergiert in einer wesentlichen Erkenntnis: Unabhängig von der spezifischen kulturellen und gesellschaftlichen Umwelt der Experimente, unabhängig von unterschiedlichen Persönlichkeitsmerkmalen und auch weitgehend unabhängig von einer Variation des experimentellen Settings muss festgestellt werden, dass viele Menschen bereit sind, andere Menschen mit teilweise gravierenden Sanktionen auch für ein marginales und harmloses „Fehlverhalten“ zu bestrafen, wenn das von einer Autorität verlangt wird.

Das Grunddesign des Experiments ist auch in diesem Fall genial einfach. Es besteht aus einem Versuchsleiter, einer Versuchsperson, die die Rolle

eines „Lehrers“ übernimmt, sowie einem angeblichen „Schüler“, der jedoch Mitglied des Teams ist, ohne dass das der Versuchsperson bekannt ist. Die Versuchspersonen sind „normale Menschen“, sie sind nach dem Zufallsprinzip ausgewählt und weisen keine besonderen Persönlichkeitsmerkmale auf. Sie erhalten für die Teilnahme an dem Experiment einen gewissen Geldbetrag, es ist jedoch klar, dass sie das Geld auch dann behalten dürfen, wenn sie das Experiment aus irgendeinem Grund vorzeitig abbrechen.

Der Schüler sitzt in einem anderen Raum wie der Lehrer, ist jedoch hörbar und durch eine Glasscheibe sichtbar. Der Versuchsleiter erklärt dem Lehrer das vorgebliche Experiment und die Aufgabe des Lehrers: Man wolle in dem Experiment untersuchen, wie sich Sanktionen auf die Lernfähigkeit auswirkten. Der Schüler müsse eine bestimmte Aufgabe lösen, sich z.B. bestimmte Vokabeln merken. Wenn er einen Fehler mache, solle der Lehrer ihn mit einem Stromstoß bestrafen. Die Stromstöße seien zunächst sehr schwach und auch wenig schmerzhaft, würden jedoch bei wiederholten Fehlern verstärkt, bis zu 450 Volt, was dann auch erheblichen Schmerz verursachen würde. Der Versuchsleiter hält sich mit Interventionen zurück und übt wenig Druck aus. Zögert die Versuchsperson, weist er etwa darauf hin, dass das Experiment bis zum Ende durchgeführt werden müsse, damit die Ergebnisse aussagekräftig sind, er droht aber nicht oder zeigt auch sonst keine besonderen Emotionen.

Die Ergebnisse der Experimente sind bekannt: In dem ursprünglichen Experiment verabreichen 65% der Versuchspersonen nach und nach Stromschläge bis zu 450 Volt, obwohl die vermeintlichen Schüler vor Schmerzen schreien und darum flehen, dass das Experiment abgebrochen wird. Außerdem vermeldet ein Warnzeichen, dass bei Stromstößen dieser Größenordnung gesundheitliche Gefahren bestehen. Immer noch 40% der Versuchspersonen gehen bis zum Ende, wenn der Schüler im gleichen Raum anwesend ist, und immerhin noch 30%, wenn sie seinen Arm mit Gewalt auf die Elektroden drücken müssen (Milgram 1974).

Die zahlreichen Varianten des Experiments ändern nichts an dem wesentlichen Ergebnis, dass ein erschreckend hoher Prozentsatz von „normalen Menschen“ bereit ist, ein triviales und unerhebliches Fehlverhalten mit gravierenden Sanktionen zu belegen, dass sie bereit sind, andere Menschen exzessiv zu bestrafen und ihnen große Schmerzen zuzufügen, wenn sie dazu von einer als legitim betrachteten Autorität aufgefordert werden – ohne selber Sanktionen oder Nachteile befürchten zu müssen, wenn sie diesen Anforderungen nicht nachkommen. Die Experimente zeigen allerdings auch, dass die meisten Versuchspersonen ihre Aufgabe nicht ohne Vorbehalte und

nicht aus eigenem Antrieb erledigen. Sie müssen wiederholt ermahnt und gedrängt werden und zeigen deutliche Symptome von Stress und Widerwillen.

Milgram selber war überrascht und erschüttert von den Ergebnissen seiner Experimente. Er folgerte aus ihnen, dass es nicht mit bestimmten abweichenden Persönlichkeitsmerkmalen oder besonderen kulturellen und gesellschaftlichen Faktoren zusammenhängt, wenn Menschen bereit sind, in rücksichtsloser Weise ihre Mitmenschen zu misshandeln und ihnen große Schmerzen zuzufügen, nur aufgrund der Tatsache, dass dieses Verhalten von einer Autorität verlangt wird. Von seinem ursprünglichen Plan, die Experimente speziell in Deutschland zu wiederholen, weil er dort aufgrund der Erfahrungen mit dem Holocaust eine besonders ausgeprägte Autoritätshörigkeit des „German Character“ vermutete, nahm Milgram daraufhin Abstand. Er war zu der Überzeugung gekommen, dass Gehorsamsbereitschaft gegenüber Autoritäten ein universell verbreitetes Phänomen darstellt, das in bestimmten Situationen generell auftritt. Anstatt spezifische Ursachen für das als abweichend angenommene Verhalten der Deutschen im Dritten Reich zu finden, fand Milgram eine allgemeine Bereitschaft von Menschen, sich auch aufgrund unbedeutender Anlässe in grausamer und menschenverachtender Weise zu verhalten und sich einer Autorität unkritisch zu unterwerfen. Es konnte kein spezifisch „autoritärer Charakter“ als Ursache für ein solches Verhalten identifiziert werden.

Im Unterschied zu den Joy of Destruction-Experimenten geht es bei den Milgram-Experimenten um Strafen und Sanktionen: Die Versuchspersonen fügen Übel zu als Reaktion auf ein angenommenes Fehlverhalten. Sie offenbaren keine bedingungslose Bereitschaft, andere zu schädigen und in ihrem Wohlergehen zu beeinträchtigen. Es machte den Versuchspersonen auch keinen „Spaß“, andere zu bestrafen: Sie haben augenscheinlich nicht aufgrund eigener Präferenzen, sondern aufgrund der Anordnung einer als Autorität betrachteten Person gestraft.

Allerdings dokumentieren die Experimente die Bereitschaft, auch bei geringfügigem und trivialem Fehlverhalten gravierende Sanktionen zu verhängen und erhebliche Schädigungen des Opfers in Kauf zu nehmen – wobei das sanktionierte Verhalten noch nicht einmal als absichtliches Handeln betrachtet werden kann, sondern als kognitives Versagen bei der Lösung einer Aufgabe. Viele Versuchspersonen zeigen geringe Hemmungen gegen Exzesse und einen unerschütterlichen Gehorsam gegenüber einer Autorität.

Auch die Milgram-Experimente unterstützen demnach prima facie die Schlussfolgerungen aus den Joy of Destruction-Experimenten: Es wäre dann

mit einer erheblichen Sanktionsbereitschaft zu rechnen, die auch vor exzessiven Strafen nicht zurückschreckt – man dürfte also keine „natürlich“ gegebenen Reserven gegen Strafen und auch nicht gegen Strafen im Übermaß erwarten, wenn sie nur als „hoheitlich“ legitimiert erscheinen. Des Weiteren erscheint auch aus der Perspektive der Milgram-Experimente das Problem der Kostspieligkeit von Strafen für den Strafenden als eine *quantité négligable*, wenn eine Vielzahl von Menschen offenbar bereit sind, solche „Kosten“ schon aufgrund der ihrerseits sanktionsfreien Anordnungen einer Autorität in Kauf zu nehmen.

Die dritte Variante von Experimenten markiert einen Übergang zu Experimenten mit deutlich freundlicheren Ergebnissen.

### 1.3 Ultimatum-Experimente

Ähnlich wie die Milgram-Experimente sind auch die Ultimatum-Experimente in vielen Varianten repliziert und abgewandelt worden. Aus ihnen hat sich mittlerweile ein veritables Forschungsprogramm entwickelt (einen Überblick gibt Fehr/Fischbacher 2005). Auf den ersten Blick zeigen diese Experimente eine sympathische Seite des Menschen, nicht zuletzt werden sie ja auch interpretiert als Widerlegung des Modells des *Homo oeconomicus* und seiner Prämisse, dass alle Menschen in allen Situationen ausschließlich eine rationale Maximierung ihres Eigennutzes betreiben. Tatsächlich geht es aber auch bei diesen Experimenten um die Offenlegung einer augenscheinlich intrinsisch motivierte Strafbereitschaft, die persönliche Kosten in Kauf nimmt, um andere Menschen für ein unerwünschtes Verhalten zu sanktionieren – allerdings unter Bedingungen, die diese Strafbereitschaft in einem durchaus positiven Licht erscheinen lassen.

Der Urtyp des Ultimatum-Experiments wurde von Werner Güth, Rolf Schmittberger und Bernd Schwarze konzipiert und erstmals Anfang der 1980er Jahre durchgeführt (Güth/Schmittberger/Schwarze 1982). Sein Design ist ebenfalls bestechend einfach. Es sind zwei Versuchspersonen involviert, die sich nicht kennen und weder vor noch nach dem Experiment einen direkten Kontakt miteinander haben. Eine der Versuchspersonen – der Proponent – erhält einen Geldbetrag, der durchaus beträchtlich sein kann. Er muss daraufhin ein Angebot an die andere Versuchsperson – den Responder – machen, ihm einen nach dem Gutdünken des Proponenten festgelegten Anteil an diesem Betrag zu überlassen. Nimmt der Responder das Angebot an, können beide die betreffenden Geldbeträge behalten. Lehnt der

Responder ab, gehen beide leer aus: Das Experiment wird jeweils nur eine Runde gespielt.

Unterstellt man, dass die Versuchspersonen sich ausschließlich und rational an ihrer Gewinnerwartung orientieren, müsste sich der Responder mit einem minimalen Angebot zufrieden geben und der Proponent ein entsprechend geringfügiges Angebot machen: Denn bei der Ablehnung eines Angebots – gleichgültig wie niedrig es ausfällt – stellt sich der Responder monetär auf jeden Fall schlechter als wenn er das Angebot annimmt. Aus der Sicht eines *Homo oeconomicus* ein Angebot, das er nicht ablehnen kann!

Das tatsächliche Verhalten der Versuchspersonen widerspricht einer solchen Prognose jedoch klar. In der ursprünglichen Variante des Experiments werden Angebote unter 25% des ausgegebenen Geldbetrags mit hoher Wahrscheinlichkeit abgelehnt. Entsprechende Erwartungen aufseiten des Proponenten aber möglicherweise auch andere Motive führen zu Angeboten, die häufig an die 50%-Marke reichen.

Wie gesagt, werden diese Ergebnisse gemeinhin interpretiert als Widerlegung eines egoistischen Menschenbilds. Andere Motive als rationale Nutzenmaximierung scheinen eine verhaltensbestimmende Rolle zu spielen: etwa eine Präferenz für Fairness oder eine Aversion gegen Ungleichheit. Die Sanktionierung „unfairer“ Proponenten durch die Responder wird von einigen Autoren auch als „altruistisches Strafen“ etikettiert, weil Sanktionen verhängt würden, die nicht dem Sanktionierenden selbst zugutekommen, sondern anderen Menschen, die von einem möglicherweise geänderten Verhalten des Proponenten in Zukunft profitieren könnten. Eine solche Sichtweise wird unterstützt durch Experimentvarianten, in denen die Proponenten sogar von unbeteiligten Dritten für als unangemessen empfundene Angebote bestraft werden (Fehr/Fischbacher 2004).

Anders als im Fall der Milgram-Experimente empfindet man das Strafverhalten der Versuchspersonen im Ultimatum-Experiment tatsächlich kaum als abstoßend oder fragwürdig. Im Gegenteil: Es handelt sich um die uneigennützig Sanktionierung von Verhaltensweisen, die gemeinhin als berechtigt anerkannte Erwartungen an Fairness und Gleichheit enttäuschen. Fairness und Gleichheit verkörpern darüber hinaus für viele Menschen zentrale Werte menschlichen Zusammenlebens, es geht nicht um die Sanktionierung trivialer Fehlleistungen ohne weitere gesellschaftliche Bedeutung. Schließlich beobachtet man keine Exzesse, sondern moderate und verhältnismäßige Reaktionen, durch die der Bestrafte im Grunde nicht schlechter gestellt wird als zuvor: Es wird ihm nur ein als unverdient empfundener Gewinn verwehrt.

Trotzdem: Auch in den Ultimatum-Experimenten wird augenscheinlich eine intrinsisch motivierte Strafbereitschaft dokumentiert, es wird gestraft, ohne dass die Strafen für den Sanktionierenden selber einen Nutzen haben können. Und das unter Inkaufnahme von teilweise signifikanten Kosten: Die Beträge, auf die in einigen der Experimente die Responder verzichten, können durchaus erheblich sein. Geht es also auch hier um eine abgeschwächte Variante eines „Strafens aus Spaß“?

Jedenfalls sehen denkbare Schlussfolgerungen aus den Ergebnissen der Ultimatum-Experimente weitaus erfreulicher aus als bei den beiden anderen Experimenten. Zeigen sie doch möglicherweise, dass viele Menschen bereit sind, sich aus eigenem Antrieb und in einer überlegten und moderaten Weise für die Durchsetzung und Sanktionierung wichtiger gesellschaftlicher Werte und Verhaltensweisen zu engagieren, auch wenn das mit fühlbaren persönlichen Kosten verbunden sein mag. Der Bedarf an Substitution durch ein staatliches Strafrecht, wenn gesellschaftliche Mechanismen der Normsicherung nicht mehr ausreichend funktionieren oder außer Kontrolle geraten, könnte aus dieser Sicht geringer ausfallen als vielleicht befürchtet.

## 2. Interpretation der Experimente

Die Experimente senden zwei gegensätzliche Botschaften aus, eine eher ungemütliche, eine eher tröstliche. Die ungemütliche Botschaft lautet: Viele Menschen offenbaren eine leicht abrufbare Bereitschaft, anderen Menschen möglicherweise gravierende Übel zuzufügen und ihnen zu schaden, sei es, weil sie selber in einer deprivierten materiellen Lage sind oder weil eine Autorität sie auffordert, andere Personen zu bestrafen. Die tröstliche Botschaft lautet: Viele Menschen sind bereit, wichtige soziale Werte wie Fairness oder Gleichheit „selbstlos“ durchzusetzen und sozial unerwünschte Verhaltensweisen, die solche Werte missachten, unter Inkaufnahme persönlicher Kosten zu sanktionieren.

Doch wie kann man diese in einem experimentellen Setting evozierten Verhaltensweisen erklären? Zugespielt gefragt: Handelt es sich bei diesen Verhaltensweisen um einen Ausdruck der Natur des Menschen oder um ein Ergebnis der Gesellschaft und Kultur? Müssen wir mit psychologischen Konstanten und tief verankerten, evolutionär entstandenen Verhaltensdispositionen rechnen, oder geht es um die soziale Einbettung von Verhalten und externe Einflussfaktoren in der Umwelt der handelnden Akteure?

### 2.1 Joy of Destruction-Experimente

Die Autoren der Joy of Destruction-Experimente legen schon auf der deskriptiven Ebene durch ihre Wortwahl eine psychologische Deutung nahe: Zerstörungsfreude, Boshaftigkeit, Neid oder Missgunst beziehen sich auf Emotionen und Gefühle und ein Verhalten, das als affektiv motivierte Reaktion auf ein Leben in Armut gedeutet werden kann. Äußere Faktoren in der natürlichen und gesellschaftlichen Umwelt der Akteure wären dann zwar die Auslöser für das beobachtete Verhalten, entscheidend aber wäre das Vorhandensein entsprechender emotionaler Dispositionen. Schon die Namensgebung für die Experimente suggeriert also eher einen psychologischen als einen sozialen Mechanismus, ein eher gefühlsgesteuertes und spontanes als ein überlegtes und kalkuliertes Verhalten.

Eine andere Interpretationshypothese ist aber möglich. Inspiration dazu kann man sich bei einer klassischen Feldstudie aus den 1950er Jahren holen: Edward Banfields „The Moral Basis of a Backward Society“ (1958). Banfield untersuchte während eines mehrmonatigen Forschungsaufenthalts in Süditalien die sozialen Beziehungen in einem rückständigen Dorf in der Provinz Potenza. Auch in diesem Dorf herrschten ärmliche Verhältnisse, harte Umweltbedingungen und ungünstige Entwicklungsperspektiven. Auch hier hätte man deshalb in einer optimistischen Sicht der Dinge erwarten können, dass die Bewohner in besonders intensiver Weise kooperieren, um durch gebündelte Kräfte und gemeinsames Handeln ihre Probleme besser zu bewältigen.

Das Gegenteil war der Fall. Banfield beobachtete einen zutiefst feindseligen und endemisch misstrauischen Umgang der Einwohner miteinander, bei dem es nur um den Nutzen der eigenen Kleinfamilie ging und man dem Nachbarn jedes Stück Wurst und jedes Hühnerei missgönnte. Banfield fasste diese Haltung als „amoralischen Familismus“ zusammen. Sie machte jede Form der Kooperation unmöglich, sowohl den bilateralen Austausch unter Nachbarn, von dem beide Parteien wechselseitig hätten profitieren können,<sup>2</sup> als auch die Realisierung gemeinsamer Projekte, wie etwa eine Verbesserung der dörflichen Infrastruktur.

Banfield erklärt dieses Verhalten aber nicht psychologisch als Ergebnis von Leidenschaften und Emotionen, sondern als Ergebnis von Verhaltens-

2 Im Sinne David Humes: „Dein Korn ist heute reif, das meinige wird es morgen sein. Es ist für uns beide vorteilhaft, dass ich heute bei Dir arbeite und Du morgen bei mir.“

prinzipien, die als implizite „strategische Empfehlungen“ Teil des amoralischen Familismus sind. Eines der zentralen Prinzipien lautet: „Jeder ist ein potentieller Feind, der geschwächt werden muss. Geht es dem Nachbarn besser, ist das für einen selbst schädlich.“ (Banfield 1958, 111)

Einem solchen Prinzip zu folgen muss keineswegs ein Zeichen von Irrationalität oder Kurzsichtigkeit sein und auch nicht notwendigerweise auf Affekte wie Boshaftigkeit oder Missgunst zurückgehen. Zwei unterschiedliche Szenarien sind denkbar. Zum einen ist es in einer Situation mit extremer Ressourcenknappheit möglich, dass tatsächlich kein Gewinn aus einem kooperativem Austausch erwartet werden kann: Wenn die Herde, die man zum Überleben braucht, ein Minimum von Weideland benötigt, das von einem anderen Herdenbesitzer ebenfalls beansprucht wird, dann kann die Schwächung des Wettbewerbers, seine Vertreibung oder zumindest die Minderung seines Viehbestandes eine rationale Überlebensstrategie sein. In einem extrem kompetitiven Umfeld, in dem es um den Zugang zu überlebenswichtigen knappen Ressourcen geht, können Nullsummenspiele entstehen, bei denen Kooperation keinen aufteilbaren Effizienzgewinn verspricht.

Doch selbst wenn Kooperation einen solchen Nutzen im Prinzip ermöglichen würde, können die Beteiligten in einem „Gleichgewicht wechselseitigen Misstrauens“ gefangen sein – das ist die Diagnose von Robert Putnam in seinen eigenen Studien zu Südtalien (Putnam 1993, 177ff.). Wenn man damit rechnen muss, dass potentielle Interaktionspartner jeden kurzfristigen Vorteil, den sie auf Kosten anderer realisieren können, ohne zu zögern wahrnehmen werden, dann wird man rationalerweise auf der Hut davor sein, sich durch Vorleistungen oder einseitiges Vertrauen verletzlich zu machen – insbesondere dann, wenn auch relativ geringfügige Nachteile und Verluste existenzbedrohend sein können. Es kann ein Teufelskreis des Misstrauens entstehen, für den ebenfalls nicht Emotionen und Affekte verantwortlich sind, sondern der tatsächlich das Ergebnis eines individuell-rationalen Handelns sein kann.

Auf diesem Hintergrund wird eine alternative Sichtweise des Verhaltens der Versuchspersonen in den Joy of Destruction-Experimenten plausibel. Anstatt auf psychologische Mechanismen rekurriert sie auf die gesellschaftliche Umwelt und die sozialen Beziehungsstrukturen der Versuchspersonen. Menschen, die in einer Welt mit extremer Ressourcenknappheit und einem endemischen Misstrauen in ihrer Gemeinschaft leben, könnten demnach zu einem zerstörerischen Verhalten neigen, weil sie von einer Schwächung oder Schädigung ihrer Wettbewerber um diese knappen Ressourcen profitieren. Ein Bezug auf Affekte oder gefühlsbestimmte Verhaltensweisen wäre dem-

nach als Erklärung nicht notwendig: Es wäre vielmehr unter den gegebenen Bedingungen ein instrumentelles und strategisches Verhalten, das jede Gelegenheit ergreift, um einen komparativen Vorteil gegenüber einem Konkurrenten zu erzielen. Gefühle wie Zerstörungslust und Schadenfreude könnten dann als Folge einer solchen Situation entstehen – sie wären aber nicht ihre Ursache.

## 2.2 Milgram-Experimente

Wenn man sich den Milgram-Experimenten und ihrer Interpretation zuwendet, ist man zunächst einmal mit den Erklärungsansätzen von Milgram selber konfrontiert. Sie erscheinen aus heutiger Sicht wenig überzeugend. Nachdem Milgram zu seinem Leidwesen akzeptieren zu müssen glaubte, dass blinde Folgebereitschaft gegenüber einer Autorität keineswegs ein Spezifikum der Deutschen unter dem Einfluss des Nationalsozialismus war, sondern eine erschreckend allgemein verbreitete Verhaltensdisposition, erklärte er Autoritätsgehorsam zu einer quasi anthropologischen Konstante mit einer neurobiologischen Grundlage. Die Gegenwart und der Einfluss einer Autorität führten demnach zu „Verschiebungen“ im Muster der nervlichen Funktionen und bewirkten eine psychologische Transformation in einen „Agens-Zustand“ der Verantwortungslosigkeit und der Eingliederung in eine Hierarchie. Personen im Agens-Zustand würden zu einem von ihrem früheren Selbst verschiedenen Wesen mit neuen Eigenschaften, die mit ihrer normalen Persönlichkeit nur wenig zu tun haben müssen (Milgram 1997). Auslöser dieses Mechanismus ist allein die „Macht der Situation“, die Präsenz einer als legitim betrachteten Autorität.

Zu einem solchen „situationistischen“, psychologischen Ansatz gibt es aber auch im Fall der Milgram-Experimente eine Alternative, die das Verhalten der Versuchspersonen auf ihre soziale Einbettung zurückführt. Die überraschend weite Verbreitung des beobachteten Verhaltens lässt sich aus dieser Perspektive mit einer entsprechend weiten Verbreitung der für dieses Verhalten bestimmenden gesellschaftlichen Institution erklären.

Bei dieser Institution handelt es sich um die Wissenschaft. In allen modernen Gesellschaften wird der Institution der Wissenschaft eine hohe Legitimität und ein bedeutender gesellschaftlicher Wert zugeschrieben. Entsprechend groß ist das Ansehen der Wissenschaftler als anerkannter und autorisierter Vertreter dieser Institution, denen in der Regel persönliche Unabhängigkeit und Integrität attestiert werden. Die Institutionalisierung der

Wissenschaft ist Ergebnis und Verkörperung der umfassenden epistemischen Arbeitsteilung zwischen Experten und Laien in allen Bereichen einer entwickelten Gesellschaft. Hat sich diese Arbeitsteilung erfolgreich und nachhaltig etabliert, mündet sie in ein belastbares Vertrauensverhältnis. Dieses Vertrauen führt dazu, dass Laien in vielen Kontexten bereit sind, den Informationen, Diagnosen und Anleitungen von wissenschaftlichen Autoritäten zu folgen, und zwar gerade auch dann, wenn sie die Wahrheit der Auskünfte, die Richtigkeit von Diagnosen und die Angemessenheit von Anleitungen nicht aufgrund eigener Urteilsfähigkeit und Kompetenz beurteilen können (Baurmann 2010). Diese Folgebereitschaft wird durch soziale Normen verstärkt. Sie fordern ein solches Vertrauen ein und verlangen ausdrücklich, dass man die überlegene Kompetenz wissenschaftlicher Autoritäten anerkennen und sich ihren Erkenntnissen und Einschätzungen beugen soll – von der Gesundheitsvorsorge bis hin zur Kindererziehung.

Aus dieser Perspektive lässt sich das Verhalten der Versuchspersonen in den Milgram-Experimenten als Folge epistemischer Arbeitsteilung, der sozialen Normen, die im Rahmen dieser Arbeitsteilung Folgebereitschaft gegenüber wissenschaftlichen Experten verlangen, sowie einem gewachsenen Vertrauen in die Kompetenz und Integrität von Wissenschaftlern erklären. Das Mantra der Wissensgesellschaft lautet: Laien können und sollten sich nicht anmaßen, die Handlungsweisen von wissenschaftlichen Autoritäten in ihrer wissenschaftlichen Forschung in Frage zu stellen! Die Versuchspersonen, die den Anordnungen der angeblichen Versuchsleiter Folge geleistet haben, hätten das nach dieser Hypothese nicht aufgrund eines blinden Gehorsams gegenüber einer beliebigen Autoritätsfigur getan, sondern aufgrund ihrer Annahme, dass sie einen authentischen Vertreter der vertrauenswürdigen und gesellschaftlich hoch geschätzten Institution Wissenschaft vor sich haben, der in eine wissenschaftliche Forschungsaufgabe involviert ist, die sie als Laien nicht umfassend verstehen können.<sup>3</sup>

Damit soll keineswegs gesagt sein, dass das Verhalten der Versuchspersonen in den Milgram-Experimenten zu begrüßen ist. Natürlich wäre es wünschenswert, dass sie nicht ein so großes Vertrauen in die Wissenschaft dokumentiert hätten, sondern in einer solchen Situation aufgrund gesunden Menschenverstandes und ihrer moralischen Intuitionen die Integrität und Vertrauenswürdigkeit des „Wissenschaftlers“ in Zweifel gezogen hätten –

insbesondere angesichts der Tatsache, dass in dem Experiment andere „Versuchspersonen“ offenbar große Schmerzen erleiden mussten.

Aber die Fragwürdigkeit betrifft vor allem die sozialen Normen, die im Hintergrund das Vertrauen in die Wissenschaft und ihre unanfechtbare Integrität einfordern, und eine soziale Praxis, die ein solches Vertrauen stützt und ein unrealistisches Image von Wissenschaftlern, ihrer Integrität und Unfehlbarkeit vermittelt. Man kann nicht behaupten, dass blindes Vertrauen und Gehorsam verlangt werden, denn Vertrauen und Folgebereitschaft werden ja an ausgefeilte Indikatoren für den Status einer wissenschaftlichen Autorität gebunden. Aber man kann einfordern, dass Vertrauen und Folgebereitschaft gegenüber wissenschaftlich legitimierten Autoritäten in stärkerem Maße relativiert werden durch realistische Erwartungen, kritische Reflexionsfähigkeit und angemessene Skepsis. In einer Gesellschaft, die in ihren Kernbereichen von der Leistungsfähigkeit und den Erkenntnissen der Wissenschaft abhängig ist, müssen auch Nicht-Wissenschaftlern und Laien die notwendigen Kompetenzen und Informationen vermittelt werden, damit sie zu einer kritischen Beurteilung der wissenschaftlichen Praxis und der Wissenschaftler in der Lage sind.

Das Fazit aus Sicht der alternativen Hypothese lautet jedenfalls: In den Milgram-Experimenten offenbart sich nicht eine anthropologisch konstante Disposition, beliebigen Autoritäten bedingungslos Gehorsam zu leisten, sondern es spiegelt sich die soziale Einbettung der Versuchspersonen in eine weit verbreitete gesellschaftliche Praxis der kognitiven Arbeitsteilung, die von Laien verlangt, der besonderen Kompetenz und persönlichen Integrität von wissenschaftlichen Experten zu vertrauen und ihren Anweisungen Folge zu leisten, auch dann, wenn sie Sinn und Zweck dieser Anweisungen nicht immer verstehen mögen.

### 2.3 Ultimatum-Experimente

Kommen wir zu der Alternative „Natur oder Kultur“, Psychologie oder soziale Einbettung bei der Interpretation der Ultimatum-Experimente. Auf den ersten Blick liegt hier eine Erklärung durch soziale Einbettung nahe, denn augenscheinlich existieren in den meisten Gesellschaften soziale Normen der Fairness und Gleichheit, die als Verhaltensdeterminanten bei diesen Experimenten in Frage kommen. Bei genauerem Hinsehen erkennt man jedoch, dass man es bei den Ultimatum-Experimenten mit zwei unterschiedlichen Verhaltensweisen zu tun hat, die separate Erklärungen erfordern.

3 Im Übrigen lagen sie mit dieser Annahme ja nicht falsch: Tatsächlich hatten sie es mit einem integren Vertreter der Wissenschaft zu tun, und tatsächlich ist den anderen „Versuchspersonen“ kein Haar gekrümmt worden.



Erstens geht es um das Angebot des Proponenten und die Bewertung dieses Angebots durch den Responder. Dabei erscheinen Prinzipien der Fairness und Gleichbehandlung relevant, wenn der Proponent einen signifikanten Anteil von seinem Geldbetrag anbietet oder wenn der Responder ein Angebot als angemessen oder unzureichend einschätzt. Zweitens geht es zusätzlich um die Bereitschaft vieler Responder, als unangemessen empfundene Angebote zurückzuweisen und damit einen Bruch von Fairness- und Gleichheitsprinzipien unter Inkaufnahme persönlicher Kosten zu sanktionieren. Es stellen sich also zwei Fragen: Warum verhalten sich Versuchspersonen entsprechend der Prinzipien der Fairness und Gleichheit? Und warum sind Versuchspersonen bereit, Verhalten, das diese Prinzipien missachtet, „selbstlos“ zu sanktionieren?

Bei der ersten Frage lautet die in der Tat naheliegende Antwort: Weil diese Prinzipien eben als soziale Normen Teil der gesellschaftlichen Praxis sind. Dennoch sind andere Hypothesen möglich. Menschen könnten auch deshalb eine genuine Präferenz für Fairness und eine Ungleichheitsaversion offenbaren, weil eine solche Disposition Teil der menschlichen Natur und das Ergebnis eines evolutionären Prozesses ist – eine solche Erklärungshypothese für eine Fairnesspräferenz findet sich z.B. bei Ken Binmore (2005).

Die Hypothese der sozialen Einbettung unterstellt dagegen, dass die Versuchspersonen in den Ultimatum-Experimenten gesellschaftliche Werte der Fairness und Gleichheit akzeptieren und entsprechende soziale Normen befolgen, wenn sie ihren Geldbesitz aufteilen oder das ihnen offerierte Angebot bewerten. Ebenso wie bei den Milgram-Experimenten kann man weiterhin argumentieren, dass die relevanten sozialen Faktoren – das Ansehen der Wissenschaften einerseits, Normen der Fairness und Gleichheit andererseits – weit verbreitet sind und in vielen Gesellschaften und Kulturen zu den etablierten Bestandteilen der gesellschaftlichen Praxis gehören. Eine Erklärung für die Replizierbarkeit der Experimente in verschiedenen Kulturräumen und Gesellschaften muss deshalb nicht auf kultur- oder gesellschafts unabhängige anthropologische Faktoren rekurren.

Die zweite Frage scheint allerdings die gerade gegenteilige Antwort nahelegen: Für die „uneigennützig“ Sanktionierung unerwünschter Verhaltensweisen, die anerkannte Prinzipien und Normen verletzen, ist eine Erklärung, die sich auf evolutionär entwickelte retributive Emotionen und Vergeltungswünsche bezieht, für viele möglicherweise plausibler als eine Erklärung, die nach einer sozialen Einbettung für ein solches Verhalten Ausschau hält. Das signalisiert man auch, wenn man etwa das Sanktionsverhalten der Versuchspersonen unter den Begriff des „altruistischen Stra-

fens“ subsumiert. In der Tat dient ja die Bestrafung von Versuchspersonen im Ultimatum-Experiment oder in ähnlichen experimentellen Settings nicht den Sanktionierenden selber, sondern verursacht ihnen nur Kosten. Der Begriff des „altruistischen Strafens“ legt nahe, dass sie mit ihren Sanktionen nicht sich selbst, sondern anderen nützen wollen, indem sie die Sanktionierten in einer Weise beeinflussen, die zukünftigen Interaktionspartnern zugutekommen könnte – und dass ihr Verhalten auf ein bestimmtes intrinsisches Motiv zurückgeht: „Altruismus“ als eine affektuelle Disposition, die Interessen anderer Menschen in die eigenen Präferenzen zu integrieren.

Aber auch in diesem Fall kann eine alternative Hypothese formuliert werden, die das offenbarte Sanktionsverhalten in den Ultimatum-Experimenten mit seiner sozialen Einbettung erklärt. Sie kann sich auf eine weitere praktisch universell verbreitete gesellschaftliche Institution beziehen: die soziale Norm der Reziprozität. Gemäß dieser Norm wird ein reziprokes Verhalten auch dann erwartet und gefordert, wenn ein solches Verhalten der handelnden Person keinen greifbaren Nutzen vermittelt: Wir sollen uns dankbar erweisen und andere Personen „belohnen“, wenn sie uns oder der Allgemeinheit einen Dienst erweisen; und wir sollen unseren Unmut ausdrücken und andere Personen „bestrafen“, wenn sie uns oder der Allgemeinheit schaden und ein sozial unerwünschtes Verhalten zeigen – gleichgültig, ob wir selber von unserer Dankbarkeit und unserem Unmut profitieren oder nicht.

Das Handeln der Versuchspersonen im Ultimatum-Experiment entspricht einer solchen Reziprozitätsnorm. Der Responder, dem ein unfaires Angebot unterbreitet wird, erlebt dabei nicht nur persönlich ein Verhalten, das er als unangemessen empfindet, sondern dieses Verhalten verletzt darüber hinaus eine weithin akzeptierte Fairnessnorm, ist also auch ein sozial unerwünschtes und damit sanktionswürdiges Verhalten. Insofern lässt sich das Handeln des Responders insgesamt sowohl unter Bezug auf soziale Normen der Fairness und Gleichheit als auch der Reziprozität erklären (Elster 2005).

Die Plausibilität dieser „Einbettungshypothese“ im Vergleich zu einer auf retributive Emotionen abzielenden „Vergeltungsthese“ lässt sich stützen, wenn man auf zwei weitere Befunde hinweist. Erstens gibt es eine signifikante Zahl von Respondern, die bereit sind, auch in ihrer absoluten Größe erhebliche Beträge zurückzuweisen, wenn die Angebote der 50%-Marke nicht nahe genug kommen. Retributive Emotionen würde man in den Ultimatum-Experimenten aber eher mit besonders kleinen Angeboten verbinden, bei denen Empörung und Ärger aufseiten der Responder unmittelbar nachvollziehbar wären. Zweitens zeigen bereits erwähnte Experimentvari-

anten, dass viele Versuchspersonen auch bereit sind, unter Inkaufnahme eigener Kosten Proponenten zu bestrafen, die anderen Personen „zu niedrige“ Angebote machen. In beiden Fällen scheint eine Hypothese plausibler, die eine Sanktionsbereitschaft eher als Folge einer als „Verpflichtung“ empfundenen sozialen Forderung erklärt denn als einen Ausdruck emotional verankerter Vergeltungsgefühle oder altruistischer Präferenzen.

Schließlich ist eine Erklärung des experimentell nachgewiesenen Sanktionsverhaltens bei den Ultimatum- aber auch den Milgram-Experimenten durch soziale Einbettung besser in der Lage, die trotz allem erhebliche Verhaltensunterschiedlichkeit des Verhaltens bei den Versuchspersonen zu erklären. Denn die Ergebnisse der Experimente sind insgesamt zwar erstaunlich robust angesichts unterschiedlicher gesellschaftlicher und kultureller Kontexte – aber in beiden Experimenten gibt es ja trotzdem erhebliche Verhaltensunterschiede innerhalb der Gruppe der Versuchspersonen selber: Immerhin 35% der „Lehrer“ in den Milgram-Experimenten verweigern sich der Autorität, nicht alle Responder sind bereit, auf ihren materiellen Vorteil zu verzichten, um einen unfairen Proponenten zu bestrafen. Dass die Folgebereitschaft gegenüber sozialen Normen in einer Population ungleich verteilt und unterschiedlich stark ausgeprägt ist, ist aber nicht erstaunlich und erscheint prima facie plausibler als eine entsprechend ungleiche Verteilung und unterschiedliche Stärke evolutionär entwickelter Dispositionen wie etwa retributiver Emotionen. Aber welche Faktoren welche Varianz erklären können, ist eine empirisch zu klärende Frage, die nicht allein durch Experimente beantwortet werden kann.

### 3. Fazit

Nach alledem lassen sich zumindest plausible Hypothesen formulieren, mit denen die experimentellen Befunde zum Sanktionsverhalten durch die soziale Einbettung des Verhaltens erklärt werden. Zugespitzt lauten sie:

1. Armut kann „zerstörerisches“ Verhalten verursachen, weil Armut zu Verdrängungswettbewerb führen kann.
2. Hierarchien können Sanktionsexzesse ermöglichen, weil soziale Normen Folgebereitschaft gegenüber legitimierte Autoritäten verlangen.
3. Der Bruch anerkannter sozialer Normen kann selbstlose Bestrafungen auslösen, weil soziale Reziprozitätsnormen die Sanktionierung sozial unerwünschten Verhaltens fordern.

Die Hypothesen zur sozialen Einbettung des experimentell evozierten Sanktionsverhaltens unterstellen freilich die Richtigkeit einer weiteren Annahme: dass viele Versuchspersonen auch in Laborexperimenten, in denen soziale Kontrolle ausgeschlossen ist und sie in einer künstlich isolierten und abgeschotteten Welt agieren, in der sie keine Konsequenzen ihres Handelns befürchten müssen oder erhoffen dürfen, dennoch eine mehr oder weniger stark ausgeprägte Neigung besitzen, die Präferenzen, Gewohnheiten, Strategien und Normorientierungen zu reproduzieren, die sie in der sozialen Realität offenbaren und praktizieren. Die Annahme, dass Versuchspersonen eine solche Übertragung ihres Sozialverhaltens in die Experimentalsituation vornehmen, ist aber nicht nur plausibel, sondern macht Experimente erst interessant und aussagekräftig.

Nehmen wir an, die Hypothesen zur sozialen Einbettung erweisen sich als zutreffend. Ist das aus strafrechtswissenschaftlicher Sicht eine gute oder eine schlechte Nachricht?

Gut wäre die Nachricht, dass das menschliche Sanktionsverhalten weitgehend plastisch ist und nicht wesentlich durch die menschliche Natur determiniert und festgelegt ist. Gesellschaftliche Lebensbedingungen, Institutionen und Normen wären dann vielmehr entscheidend für die Art und Weise, wie wir Strafen und Sanktionen praktizieren. Und wir hätten einen großen Spielraum, wie wir diese Bedingungen und Institutionen gestalten. Also stünden auch der nachhaltigen Verankerung eines rechtsstaatlich gezähmten Strafrechts keine Restriktionen und Hürden im Wege, die sich aus psychologischen Mechanismen oder anderen anthropologischen Konstanten ergeben.

Schlecht wäre die Nachricht, dass die soziale Einbettung eines plastischen menschlichen Sanktionsverhaltens auch zu exzessiven Strafen und Schädigungen führen kann. Die menschliche Natur und Psychologie würden dann das menschliche Sanktionsverhalten zwar nicht programmieren, setzten ihm aber auch keine Grenzen der Verhältnismäßigkeit und des Mitleids – eine Befürchtung, die leider häufig bestätigt wurde.

Wenn uns demnach die Art und Weise, wie wir unsere gesellschaftliche Praxis des Strafens – und Belohnens! – gestalten, nicht durch unsere Natur vorgeschrieben wird, dann bleibt für sozialwissenschaftliche und rechtswissenschaftliche Normspezialisten wie Klaus Lüderssen noch viel zu tun!

Literatur

- Banfield, E. C. (1958): *The Moral Basis of a Backward Society*. The Free Press, New York.
- Baurmann, Michael (2010): Kollektives Wissen und epistemisches Vertrauen: der Ansatz der Sozialen Erkenntnistheorie. Soziologische Theorie kontrovers. Sonderheft 50 der Kölner Zeitschrift für Soziologie und Sozialpsychologie, S. 185-201.
- Binmore, K. (2005): *Natural Justice*. Oxford University Press, Oxford.
- Elster, J. (2005), Fehr on Altruism, Emotion, and Norms. *Analyse & Kritik* 27, S. 197-210.
- Güth, W./R. Schmittberger/B. Schwarze (1982): An Experimental Analysis of Ultimatum Bargaining. *Journal of Economic Behavior & Organization* 3, S. 367-388.
- Fehr, E./U. Fischbacher (2004): Third-Party Punishment and Social Norms. *Evolution and Human Behavior* 25, S. 63-87.
- Fehr, E./U. Fischbacher (2005): Human Altruism – Proximate Patterns and Evolutionary Origins. *Analyse & Kritik* 27, S. 6-47.
- Milgram, S. (1974): *Obedience to Authority: an Experimental View*. Harper, New York.
- Ostrom, E. (1990), *Governing the Commons. The Evolution of Institutions for Collective Action*. Cambridge University Press, Cambridge.
- Prediger, S./B. Volland/B. Herrmann (2013): *Resource Scarcity, Spite and Cooperation*. GIGA Working Paper, N. 227, June
- Putnam, Robert D. (1993): *Making Democracy Work: Civic Traditions in Modern Italy*. Princeton University Press, Princeton.

Diskussion zum Referat von Michael Baurmann

Dirk Fabricius

Ich bin sehr dankbar, besonders für das erste Experiment, das ich nicht kannte. Was das dritte Experiment angeht, könnte man ihm entnehmen, dass Strafen in dieser einmaligen Situation in der Perspektive der Evolution der Kooperation eigentlich eine Investition in die Zukunft (im Hinblick auf eine mögliche zukünftige Begegnung) ist, dass es sich also um ein iteratives Gefangenendilemma handelt. In diesem Sinn könnte man daraus schließen, dass diese auf Reziprozität angelegten Verhaltensprogramme evolutionär entstanden sind.

Die Hypothese „joy of destruction“ kommt mir dagegen zweifelhaft vor, wenn ich das Milgram-Experiment in seinen Varianten betrachte. Ich fand diese Varianten (z.B. „Der Versuchsleiter verlässt den Raum“ oder „Zwei Versuchsleiter sind da und geraten in Streit, ob man das Experiment fortsetzen sollte“) immer besonders spannend. Wenn man der „joy of destruction“-Hypothese folgen würde, müsste man wohl vorhersagen, die Leute würden in jedem Fall weitermachen und erst recht vielleicht bis zu den 450 Volt gehen. Das Erstaunliche ist aber, dass das Gegenteil der Fall ist: Es kippt total um; die Quote derjenigen, die weitermachen, geht ganz nach unten. Und bei denen, die weitermachen, findet man eine Korrelation mit dem, was man früher als Psychopathie bezeichnet hat. Das würde dann für das erste Experiment für mich Probleme aufwerfen. Was zeigt es uns eigentlich? Man müsste wohl fragen: Was würden denn diese Versuchspersonen als Gründe angeben, warum sie so handeln? Dann würde man vielleicht auch ein Stück weiterkommen. Ist tatsächlich der kleine Vorteil, der sich in so einem Rennen ergibt, wenn alle arm sind, so angsterregend, dass man befürchtet, endgültig in Todesnähe oder Todesgefahr zu geraten, weil der andere einen Vorsprung in den Ressourcen hat?

Das bringt mich zu einem letzten Punkt, der schon bei dem Vortrag von Herrn Günther, bei dem ja Breivik vorkam, anklang. Wenn man sich das Welt- und Menschenbild von Breivik vorstellt, dann hat er die Vorstellung, er befindet sich in einer Welt, in der das Volk bedroht ist. Er malt sich eine Fiktion, eine Wahnvorstellung oder eine Wahnwelt aus, um damit indirekt zu sagen: ich befinde mich in einer Notwehr- oder Notstandssituation. Und das liefert ihm dann die Basis dafür, mit dieser Zerstörungswut, mit diesem

Zerstörungsvergnügen durch die Welt zu marodieren. Für mich stellt sich dann die Frage: liegt dem tatsächlich so ein Zerstörungsvergnügen zugrunde und was werden dann für Rechtfertigungen jeweils dafür erdacht? Oder gibt es, was die Interpretation der Milgram-Experimente nahelegt, ein Element von Sadismus und Schmerzzufügung, das sich nur auf den Befehl der Autorität stützt, aber eben nicht auf ein solches primäres Vergnügen.

*Michael Baurmann*

Ja, vielen Dank. Sie erwähnen zahlreiche interessante Aspekte. Ich fange bei dem letzten Punkt an. Ich glaube, die Milgram-Experimente dokumentieren kein Strafen aus Spaß, kein Vergnügen an Schmerzzufügung, es sind keine Joy of Destruction-Experimente. Bei den Milgram-Experimenten ist es vielmehr so, dass es den meisten Versuchspersonen unangenehm ist, den „Schülern“ Schmerz zuzufügen. Man braucht allerdings auch nicht viel Druck, um sie sozusagen bei der Stange zu halten. Aber in ihren Reaktionen offenbaren sie doch: das ist nicht das, was sie eigentlich gerne machen. Bei widerstrebenden Autoritäten bricht das Ganze deshalb auch schnell zusammen. Um den gewissen Widerwillen der Versuchspersonen zu überwinden, benötigt man die Präsenz und Eindeutigkeit der Autorität, die die Institution Wissenschaft in einem optimalen Sinne repräsentiert. Es gibt ja Varianten, die sich in einem Hinterhof abspielen, der Wissenschaftler hat Flecken auf dem Kittel. Auch dann ist die Folgebereitschaft sehr viel geringer. Das zeigt, dass die Versuchspersonen nicht aus eigenem Antrieb andere verletzen oder schädigen wollen. Aber man kann sie relativ leicht dazu bringen, sobald sie das Gefühl haben, da ist wirklich ein kompetenter und glaubwürdiger Repräsentant der Wissenschaft. Und dieses Verhalten ist auch nicht so abwegig: Wenn Menschen in einer funktionierenden Wissensgesellschaft leben, in der wissenschaftlichen Institutionen eine hohe Legitimität zugeschrieben wird und Wissenschaftler ein hohes Ansehen genießen, und diese Autoritäten, die sonst nicht in Frage gestellt werden, nun Dinge verlangen, die man als Laie sich vielleicht nicht zu beurteilen traut, ist es nicht ein Zeichen für eine abnormale Persönlichkeit, wenn man sich sagt: „Na ja, das wird schon seinen guten Sinn haben, und ich sollte erst einmal machen, was von mir verlangt wird.“

Zu der Frage, ob in den Ultimatum-Experimenten das Strafen nicht als eine Investition in die Zukunft gedeutet werden kann: Der Witz der Ultimatum-Experimente ist gerade, dass die Versuchspersonen genau erkennen können, es gibt keine Iteration, ich sehe den anderen nie wieder, er wird auch

nie erfahren, wer ich bin. Mit diesem Design sollen Kalkulationen mit wiederholten Interaktionen und Lerneffekte systematisch ausgeschlossen werden. Es gibt begleitende Untersuchungen, die zeigen, dass die Versuchspersonen sich darüber auch sehr genau im Klaren sind. Sie wissen, dass das „One shot“-Situationen sind, dass man deshalb keine zukünftigen Vor- oder Nachteile aus seinem Verhalten erwarten kann. Allerdings gibt es Interpretationen der Experimente, die zu bedenken geben, dass die Versuchspersonen dem experimentellen Design und der Wahrung der Anonymität möglicherweise misstrauen und deshalb ein sozial erwünschtes Verhalten zeigen. In diesem Punkt gibt es also auch gewisse Kontroversen. Ich neige aber zu der Sichtweise, dass diese Experimente tatsächlich zeigen, dass Menschen in bestimmten Situationen bereit sind, reale persönliche Kosten in Kauf zu nehmen, um ein als unangemessen empfundenenes Verhalten mit Sanktionen zu belegen. Und sie tun das nicht, weil sie sich davon einen Vorteil erhoffen oder ihrerseits negative Reaktionen befürchten. Sie handeln so, weil sie so motiviert sind. Und die Frage ist, warum sind sie so motiviert.

Und dann noch einmal zurück zu Namibia und den Joy of Destruction-Experimenten. Ich bin ja gerade zurückhaltend und würde nicht vorschnell von Vergnügen oder Freude an Zerstörung und Übelszufügung als Motiven sprechen. Ich denke, man sollte eine alternative Erklärungshypothese ernstnehmen, die zugespitzt lautet: Es kann vollkommen rational sein, dass man unter Bedingungen extremer Ressourcenknappheit versucht, sich Ressourcen des Nachbarn anzueignen oder zumindest zu verhindern, dass er Mittel zur Verfügung hat, mit denen er im Wettbewerb um diese knappe Ressourcen gefährlich werden kann. Und wenn erst einmal eine Situation besteht, in der sich eine solche Konkurrenzsituation und ein wechselseitiges Misstrauen verfestigt haben, dann wird man schlecht beraten sein, einseitig auf das Wohlwollen der anderen zu setzen. Das war auch die Diagnose von Putnam für die Situation in Süditalien. Freilich heißt das nicht, dass es unter Bedingungen von Armut und Ressourcenknappheit immer zu einem Vernichtungswettbewerb kommen muss. Aber hier gibt es Pfadabhängigkeiten, und wenn sich eine Gesellschaft erst einmal auf dem Entwicklungspfad zu einem Gleichgewicht der Kooperationsverweigerung und des wechselseitigen Misstrauens befindet, dann ist es sehr schwer, einen anderen Weg einzuschlagen. In einer solchen Situation können dann natürlich auch Emotionen wie Feindseligkeit oder Schadenfreude entstehen – sie wären dann aber Folgen, und nicht Ursachen.

Thomas Fischer

Ich möchte auch eine Anmerkung machen zu dem ersten Experiment, weil mir das ein bisschen überinterpretiert erscheint. Ich finde es nicht so schwierig und auch nicht besonders überraschend. Es fehlt daran eine Bedingung, die, glaube ich, die Erklärung erleichtern würde: Wie stark haben sich eigentlich die beiden Reichtums- oder Armutsbedingungen unterschieden? Anders ausgedrückt. Wenn Sie 100 Strafrechtsprofessoren anonym fragen, ob anderen Strafrechtsprofessoren ein Unglück geschehen möge, und alternativ die Frage stellen, ob den Sekretärinnen der anderen Strafrechtsprofessoren ein Unglück geschehen möge, wird sich eine signifikant hohe Zahl für die erste Variante entscheiden. Da bin ich mir ziemlich sicher.

[Prof. Baurmann (Zwischenruf): Ein Verdrängungswettbewerb in einem hoch kompetitiven Umfeld.]

Das bedeutet, dass die Hypothese „Armut schafft Bosheit“, ob nun anthropologisch oder sozial, nicht stimmt. Sondern Konkurrenz schafft Bosheit. Das ist schichtenunspezifisch. Das heißt, es kann sich auf jeder beliebigen Ebene einstellen. Der Mensch differenziert sich von dem, was ihm nah und bedrohlich erscheint. Sie können das auch umdrehen: Wir finden bei unseren Armen, z. B. bei unseren Harz IV-Empfängern, eine ganz ungewöhnliche, völlig irrationale Großzügigkeit gegenüber den ganz Reichen. Denen wünscht man, dass sie noch reicher werden, noch schöner sind, noch goldenere Kronen haben. Zugleich hat man einen starken Beißreflex gegenüber Menschen aus dem unmittelbaren Umfeld. Und umgekehrt, die Reichen, oder viele von ihnen, sind sehr großzügig gegenüber Personen, die keinerlei Konkurrenz darstellen, die Ärmsten der Armen usw.; da blüht das Mitleid. Unbekannten gegenüber sind wir großzügig. Menschen, die auch nur geringfügig eine Konkurrenz darstellen könnten im sozialen Umfeld, werden dagegen bekämpft. Diese Tendenz ist meiner Ansicht nach schichtenunspezifisch. Ich glaube, abschließend, dass alle drei Beispiele gleichermaßen rational erklärbar sind und auch zusammenpassen könnten. Was man daraus ableiten könnte für die Rationalität oder Irrationalität von Strafen, vermag ich spontan nicht zu sagen. Ich halte das erste Experiment für überinterpretiert und teilweise auch für unzutreffend entfernt angesiedelt. Ich glaube, das Ergebnis stammt nicht aus Namibia, sondern aus der Mitte unserer Gesellschaft.

Michael Baurmann

Eine interessante These, aber ich glaube, am Ende doch nicht wirklich überzeugend als Interpretation der Experimente. Zunächst ist natürlich richtig, Wettbewerbssituationen erzeugen Nullsummenspiele, egal in welchem gesellschaftlichen Kontext: Die Schädigung meines Konkurrenten ist nützlich für mich, Gewinne meines Konkurrenten sind meine Verluste. Anders als bei einem Partner in einem kooperativen Austausch ist das Wohlergehen eines Konkurrenten für mich nicht etwas, wovon ich selber profitieren kann. Insofern müsste man in der Tat erwarten, dass Joy of Destruction-Experimente in allen kompetitiv geprägten Kontexten die gleichen Verhaltensweisen evozieren müssten – also auch bei Universitätsprofessoren als Versuchspersonen. Wir müssen aber die Ergebnisse der Experimente in Namibia ernstnehmen: Die „Zerstörungswut“ der Versuchspersonen aus der ärmeren Gegend war signifikant größer als derjenigen aus der relativ wohlhabenden Gegend. Diese Unterschiede müssen erklärt werden, und sie widersprechen der Annahme, dass es sich auch um Experimente aus der Mitte „unserer Gesellschaft“ handeln könnte. Ich denke, wir müssen beachten, dass unter Bedingungen relativen Wohlstands Wettbewerb und Konkurrenz durch Normen und Regeln reguliert und limitiert werden. Es handelt sich eben nicht um einen Vernichtungswettbewerb, sondern, wie Adam Smith schon betont hat, der Marktwettbewerb ähnelt dem Sportwettkampf, der nach Prinzipien der Fairness organisiert ist. Das funktioniert, wenn die Institution eines fairen Wettbewerbs auch im Interesse der Wettbewerber selber ist, wenn sie insgesamt davon profitieren, dass sie sich durch die Regeln eines solchen Wettbewerbs einschränken. Doch selbst wenn das der Fall ist, können die Anreize unterschiedlich groß sein, bei günstigen Gelegenheiten gegen diese Regeln zu verstoßen und sich auf unfaire Weise Sondervorteile gegenüber einem Konkurrenten zu verschaffen. Eine Hypothese zur Erklärung der Unterschiede im „Zerstörungsverhalten“ in den beiden Regionen in Namibia könnte demnach lauten, dass es bei extremer Armut und Ressourcenknappheit möglicherweise überlebensrelevant ist, wenn man in bestimmten Situationen einen komparativen Vorteil gegenüber einem direkte Konkurrenten erzielen kann – der Anreiz und die Disposition zur „Unfairness“ könnten entsprechend stärker ausgeprägt sein.

Das Fazit aus den Joy of Destruction-Experimenten kann deshalb nicht lauten: „Aha, aber das ist doch überall so, wo wir Konkurrenz haben.“ Sondern die Experimente werfen die Frage auf, unter welchen Bedingungen Wettbewerb zu einem Vernichtungswettbewerb mit entsprechenden „Zer-

störungsstrategien“ führt bzw. wie wir es geschafft haben, den Wettbewerb in unseren Gesellschaften so zu zähmen, dass genau das nicht passiert, was in Namibia passiert oder was Banfield aus dem von ihm studierten süditalienischen Dorf berichtet.

*Klaus Günther*

Ich habe nur eine kurze Nachfrage zum ersten Experiment. Da würde mich interessieren, was Ostrom zu diesem Experiment gesagt hat. Meines Erachtens, so laienhaft ich jetzt informiert bin, trifft es ja Ihre Hypothese nicht ganz. Denn die lautet ja, dass die „tragedy of the commons“ nicht notwendigerweise eintreten muss; nämlich dann nicht, wenn die Beteiligten anfangen, ihre Ressourcen gemeinsam zu verwalten, selbst sich an der Schaffung von Regeln zu beteiligen, nach denen die Ressourcen von jedem ausgebeutet werden dürfen. Dafür scheint mir aber eine Bedingung zu sein, dass die Beteiligten auch tatsächlich Nachbarn sein müssen [Baurmann (Zwischenantwort): Es sind Nachbarn]. Ich meine, das schließt ja dann die Möglichkeit ein, dass man auf längere Zeit einander wieder begegnet und man gerade nicht die „one shot“-Situation hat, in der man darauf vertrauen kann, dem anderen nicht wieder zu begegnen und sich daher folgenlos auf dessen Kosten Vorteile verschafft. Möglicherweise betrifft das verschiedene Situationen und das wäre dann kein richtiger Einwand; vielleicht habe ich das auch falsch verstanden.

*Michael Baurmann*

Das ist ein wichtiger Punkt. Das Namibia-Experiment stellt den Ansatz von Ostrom sicherlich nicht grundsätzlich in Frage. Eine kooperative Selbstverwaltung gemeinsamer Ressourcen unter Armutbedingungen ist auch aus ihrer Sicht keineswegs selbstverständlich. Ein Hauptbestandteil ihrer Arbeit besteht vielmehr darin, die einzelnen Faktoren zu identifizieren, unter denen eine solche Selbstverwaltung erfolgreich funktionieren bzw. unter denen sie nicht funktionieren kann. Das Vorhandensein von Misstrauen, Missgunst und Schadenfreude in einer sozialen Gruppe könnte insofern als ein – vielleicht extremes – Beispiel für eine Situation abgebucht werden, in der Kooperation und Selbstverwaltung nur schwer initiiert werden können. Ostrom würde sicherlich nicht abstreiten, dass solche Situationen möglich sind. Bei der Konferenz zu ihren Ehren in Bielefeld, auf der die Experimente aus Namibia vorgestellt wurden, war sie leider nicht anwesend, weil sie damals bereits erkrankt war.

Was die Anonymität betrifft: Natürlich wird ein asoziales Verhalten durch das experimentelle Design erleichtert, weil eine Kenntnis der anderen Person sowie wiederholte Interaktionen ausgeschlossen sind. Und nach Berichten der beteiligten Wissenschaftler ist es offenbar der Fall, dass analoge Verhaltensweisen in der sozialen Realität der Versuchspersonen auch in der ärmeren Region eher nicht zu beobachten sind. Dann stellt sich aber umso mehr die Frage, warum die Versuchspersonen ihre „Nachbarn“ so schlecht behandeln, sobald sie sich durch Anonymität sicher fühlen. Im Übrigen ist auch offener und kontinuierlicher Kontakt in einer sozialen Gemeinschaft noch keineswegs eine Garantie für vertrauensvolle Kooperation und Zusammenarbeit. Hier kann man erneut auf die klassischen Studien von Putnam und Banfield in Süditalien verweisen.

*Luigi Lombardi Vallauri*

Wir alle wissen, dass die Behandlung der Tiere, die wir essen, grauenhaft ist. Hier geht es zwar nicht um Spaß am *Strafen*. Es geht aber um Spaß am Essen, was auch enorme Schmerzen und Demütigung der Tiere impliziert. Im Experiment 2 war die legitimierende Instanz die Wissenschaft, die Autorität der Wissenschaft. Welche ist hier die legitimierende Instanz?

[Prof. Baurmann (Zwischenruf): Beim Essen?]

Beim Fleisch- und Fischessen.

*Michael Baurmann*

Das ist eine andere Situation. Wir fügen Tieren Schmerz und Leid zu, aber wir haben dadurch insgesamt keine Nachteile, im Gegenteil, wir profitieren davon. Wir genießen einen Vorteil aus der Tatsache, dass wir dieses Leid zufügen. Das ist also eine andere Grundkonstellation, die aber ebenfalls ethisch fragwürdig ist. Ich denke, wir sehen hier in einigen Kulturen einen Konflikt zwischen moralischen und sozialen Normen. Unsere sozialen Normen erlauben und rechtfertigen die Verwertung von Tieren, gleichzeitig werden ernstzunehmende moralische Normen vertreten, die reklamieren, dass die Sozialnormen in diesem Bereich falsch sind – denken Sie an Peter Singer und das Animal Rights Movement.

[Prof. Lombardi Vallauri (Zwischenruf): Ist dies ein Fall der sozialen Einbettung?]

Ja, natürlich.

*Luigi Lombardi Vallauri*

Die Ethnologie, die Moral, die Wissenschaft sagen: es ist sehr grausam. Aber wir haben nicht direkt Spaß an der Grausamkeit, aber wir nehmen sie in Kauf.

*Michael Baurmann*

So sehe ich das auch. Die Praxis, wie und was wir essen, ist dabei sozial eingebettet. Und es ist eine stark normierte soziale Praxis, bei der zahlreiche Normen eine Rolle spielen. Zusätzlich gibt es die moralische Debatte, die selber nicht eindeutig ist und sehr kontrovers. Es wird ja etwa auch die Position vertreten: Wenn die Tiere alle artgerecht und nach ökologischen Kriterien gehalten werden, dann tut man den Tieren vielleicht sogar einen Gefallen, wenn man sie isst, weil sie ansonsten gar nicht auf der Welt wären. So können sie wenigstens ein paar Jahre ihr Leben genießen und, gut, nachher werden sie halt verspeist. Jedenfalls ist unsere kulinarische Praxis bestimmt durch ein komplexes Geflecht von Sozialnormen und moralischen Normen, die insgesamt kein konsistentes Bild ergeben. Dem entspricht auch die Tatsache, dass einige Menschen kein Fleisch essen, die Mehrheit isst es nach wie vor. Insgesamt sicherlich ein exemplarisches Beispiel für die soziale Einbettung unseres Verhaltens und unserer Einstellungen.

*Cornelius Prittwitz*

Mir kommt ein Faktor besonders wichtig vor, den Du [Michael Baurmann] nicht benannt hast: nämlich die Tatsache, dass mindestens zwei der drei Situationen, eigentlich sogar alle drei, außerordentlich ungewöhnliche Situationen sind. Im Alltagsleben kommen diese Situationen so nicht vor; es sind typische Experiment-Situationen. Ich habe die Arbeitshypothese, dass Menschen in Experiment-Situationen sich fragen: „Worauf mögen die hinauswollen? Wo ist der Trick? Da stimmt doch irgendetwas nicht“. Und deswegen bewegen sich diese Menschen auf einer anderen Ebene von Rationalität oder Gründen, weil sie das spielerische Element, das experimentelle Element der Situationen erkennen. Und es ist notwendig bei allen diesen Experimenten so, dass der Experimentcharakter der Situation den Teilnehmern präsent ist; es sei denn bei den unethischen Experimenten, bei denen das Objekt nicht weiß, dass es Objekt ist. Das könnte meines Erachtens einiges erklären. Man weiß nicht genau, wie das Spiel geht und worauf es hinausläuft; es wird zwar erklärt, aber man hat eben das Misstrauen, dass

der Erklärende vielleicht doch noch etwas anderes im Sinn hat. Und mindestens bei Milgram hat sich ja bewahrt, dass die eine Sache erklärt und die andere Sache überprüft wurde. Das passt natürlich zu der von Dir erwähnten „Balance des Misstrauens“: „Trau dem Frieden nicht!“, „Sachen, die scheinbar für dich nur positiv sind, haben bestimmt einen Haken!“. Und wenn dann ungewohnte oder unerwünschte Ergebnisse kommen, dann will man als Teilnehmer nicht als „Blöder“ dastehen, um in Wirklichkeit natürlich erst recht als „Blöder“ dazustehen.

Ein zweiter Punkt: Dein Vortrag legt eine Tendenz nahe, – vielleicht stärker so, Herr Fischer, wenn wir uns (sozial oder geografisch) näher sind, als wenn wir weit voneinander entfernt sind, – den Mitmenschen und Konkurrenten kleinzuhalten. Wenn das – worauf auch immer gegründet – richtig ist, würde das den Thesen kultureller Evolution widersprechen, die besagen: der entscheidende Fortschritt des Menschen sei, dass er tatsächlich *nicht* so denkt, sondern den gemeinsamen Nutzen erkennt und eigentlich kooperativ angelegt ist. Dann wäre die Frage, ob die Reaktionen der tatsächlichen Personen im Experiment eben die Reaktionen zivilisatorisch „verdorbener“ Menschen sind, Menschen mit Erfahrungen, die sich nicht so verhalten, wie sie ursprünglich angelegt sind. Dafür würden die vielfältige Erfahrung von Druck- und Konkurrenzsituationen sprechen. Aber auch der kulturelle Kontext in einem weiteren Sinn. Ich könnte mir durchaus vorstellen, dass die Experimente vor 100 Jahren anders ausgefallen wären als heute, in einer Zeit, in der man Dir ständig suggeriert „Alles, was der andere kriegt, kriegt du nicht“.

*Michael Baurmann*

Die Frage, inwieweit die Ergebnisse von Experimenten extern valide sind, stellt sich natürlich immer. Aber auch wenn wir der Meinung sind, dass bestimmte Experimente sich manifest von realweltlichen Kontexten unterscheiden, stehen wir vor der Aufgabe, das Verhalten der Versuchspersonen zu erklären. Nicht selten können wir dann beobachten, dass auch in Experimenten, in denen von aller sozialen Einbettung abstrahiert wird und extrem künstliche Situationen erzeugt werden, die Versuchspersonen dennoch so handeln, als ob sie den sonst üblichen Einflussfaktoren – wie z.B. sozialer Kontrolle oder sozialen Normen – ausgesetzt sind. Bei der Beurteilung des wissenschaftlichen Werts von Experimenten in den Sozialwissenschaften muss man außerdem berücksichtigen, dass Experimente unterschiedliche Funktionen haben können. Sie können strikt generell formulierte Annahmen

widerlegen: So falsifiziert das Verhalten der Versuchspersonen im Ultimatum-Experiment die Annahme, dass Menschen in jeder Handlungssituation als rationale Nutzenmaximierer agieren. Experimente können neue und innovative Hypothesen generieren: So haben die Milgram-Experimente wichtige Impulse für die Holocaust-Forschung geliefert. Die Joy of Destruction-Experimente können schließlich Aufmerksamkeit auf das Phänomen asozialen Verhaltens und seiner Entstehungsbedingungen lenken. Am Beispiel der Ultimatum-Experimente kann man im Übrigen sehen, dass eine Übertragung auf realweltliche Kontexte nicht unbedingt eine große Kluft überwinden muss. So haben Verhandlungssituationen etwa beim Abschluss von Arbeitsverträgen eine ganz ähnliche Anreizstruktur und die Ergebnisse sind durchaus vergleichbar: Fairnesspräferenzen können dazu führen, dass Angebote abgelehnt werden, auch wenn dadurch persönliche Kosten entstehen.

Ist der Mensch als Ergebnis der Evolution auf Kooperation angelegt? Das würde voraussetzen, dass Kooperation in jeder Umwelt vorteilhafter ist als eine Konfliktstrategie. Davon kann man nicht ausgehen. Kooperieren sollte man nur, wenn Kooperation einen Effizienzgewinn verspricht und man auf die Kooperation der anderen setzen kann. Beide Bedingungen sind nicht immer gegeben. Evolutionäre Vorteile scheint insofern nur eine Bereitschaft zu einer bedingten Kooperation zu versprechen. Und ich habe ja die Hypothese zur Diskussion gestellt, dass das Verhalten der Versuchspersonen in den Joy of Destruction-Experimenten möglicherweise damit zu erklären ist, dass kooperatives Verhalten unter den gegebenen Lebensbedingungen vielleicht nicht ohne weiteres die rationale Strategie ist.

#### *Klaus Lüderssen*

Ich möchte nichts Kritisches sagen, sondern nur etwas ergänzen: Die Kunst als Erkenntnisquelle. Ein paar verstreute Beispiele, die mir gerade einfallen. In dem Film „Die Vergessenen“ von Buñuel werden Motivation und Handlungen der Ärmsten so anschaulich, wie das eine nach wissenschaftlichen Maßstäben abgefasste empirische Studie nicht zu leisten vermag. Ein anderes Beispiel ist der Film von Saura „Die Jagd“. Hier werden verborgene Konflikte unter den mittelständischen (wie die Soziologie sagen würde) Teilnehmern offenbar, die eine wissenschaftliche Studie niemals enthüllen könnte. Das gleiche gilt – ein drittes Beispiel, hier geht es um Buch und Film – für Carlo Levi, „Christus kam nur bis Eboli“: Heidnische Zustände unter der Bevölkerung, durch Befragung oder „Feldforschung“ nicht eruierbar.

Vielleicht kann man doch sagen, dass solche künstlerischen Darstellungen die Bedeutung von Experimenten relativieren.

#### *Michael Baurmann*

Ich widerspreche nicht, sondern möchte ebenfalls ergänzen. Sicherlich muss man den Wert von Experimenten relativieren, sie sind nicht die allein seligmachende Methode. Aus meiner Sicht ist eine der wichtigsten Funktionen von Experimenten, dass sie Maschinen für Hypothesengenerierung sind. Aber solche Hypothesen müssen anschließend mit anderen Methoden überprüft werden: Lässt sich beispielsweise eine Präferenz für Fairness auch bei tatsächlichen Tarifverhandlungen oder Hauskäufen feststellen? Ebenso können Gedankenexperimente und wilde theoretische Spekulationen Hypothesen generieren. Und natürlich kann uns auch die Kunst zu neuen und innovativen Hypothesen inspirieren. Entscheidend ist aber, dass wir in der Wissenschaft über bestimmte Verfahren und Methoden verfügen, mit denen wir diese Hypothesen intersubjektiv nachvollziehbar überprüfen können. Das heißt, ein Film kann uns auf Hypothesen bringen, auf die wir sonst nicht gekommen wären, das ist vollkommen richtig. Aber wir können uns nicht nur den Film ansehen.

#### *Klaus Lüderssen*

Ich bitte um Entschuldigung, dass ich mich noch einmal melde und betone, dass die Literatur die Armutsexperimente stark bestätigt. Herr Fischer wundert sich, warum es bei den Armen weniger Kooperationsbereitschaft gibt. Dass dies bei den Armen passiert, das ist doch unsere große negative Überraschung, damit hätte man nicht gerechnet. Das würde dann wieder kompliziert, wenn man die Praxis und Theorie des Nobelpreisträgers Julius hinzunimmt, mit den Kleinkrediten in den armen Ländern. Ich weiß nicht, ob Du das kennst? [Zwischenantwort Baurmann: Ja, klar] Da hat es geklappt. Da ist also unter den Armen die Kooperation nicht negativ. Also es bleibt eben doch sehr offen. Ich möchte noch einmal unterstreichen: Die Überraschung mit der Armut, die bestätigt die schöne Literatur weitläufig.

#### *Michael Baurmann*

Nur ein Satz dazu. Wenn meine Erklärungshypothese für die Joy of Destruction-Experimente zutrifft, dann ist das Verhalten der Versuchspersonen eben nicht damit zu erklären, dass sie intellektuell minderbemittelt oder ungebildet sind, dass sie irrational agieren oder dass emotional gesteuerte psy-



chologische Mechanismen wirken. Sondern ihr Verhalten wäre eine rational nachvollziehbare Folge der Bedingungen, unter denen sie leben. Die sind, muss man dann feststellen, nicht einfach dumm oder von Leidenenschaften getrieben. Ihr Verhalten entspricht in ihrer Situation möglicherweise einer durchaus rationalen Strategie. Umso schlimmer eigentlich.

*Lorenz Schulz*

Es heißt in der Überschrift „Strafen aus Spaß? Experimente etc.“. Diese Frage muss man ja wohl nach der Diskussion und Deinem Vortrag schlicht verneinen. Es wäre anders, wenn man sagen würde „Strafen aus Rationalität“. Da würden viele sagen „Ja“. Manche würden vielleicht an Irrationalität denken. Aber wie wäre es, wenn man „Joy of Destruction“ ins Deutsche richtig übersetzt. „joy“, das ist ja nicht Spaß, das ist Freude, und noch richtiger: Lust. Dann, meine ich, wird die Frage wirklich interessant.

*Michael Baurmann*

Durch mein Interpretationsangebot möchte ich in der Tat in Frage stellen, dass wir es mit Strafen aus Spaß zu tun haben. Nach meinen Hypothesen geht es in den Experimenten um Strafen bzw. Schädigungen aufgrund rationaler Strategien und aufgrund der Wirksamkeit von sozialen Normen, die Folgebereitschaft gegenüber gesellschaftlich legitimierten Autoritäten einfordern oder die Sanktionierung und Durchsetzung von Fairnessprinzipien verlangen. Es geht also um die soziale Einbettung des Verhaltens der Versuchspersonen.

Du hast Recht, dass eine textnahe Übersetzung von „Joy of Destruction“ eher „Lust an der Zerstörung“ lauten müsste. Doch noch einmal kurz zur Sache selbst: Ich würde nicht bestreiten wollen, dass im menschlichen Sanktionsverhalten tatsächlich anthropologische Konstanten eine Rolle spielen. Gewisse retributive Emotionen sind evolutionär sehr plausibel zu erklären. Die Frage ist nur, wie relevant sind diese Emotionen, wie viel Varianz des Sanktionsverhaltens können wir mit ihnen erklären. Da bin ich eher skeptisch. Ich bin ja auch Soziologe und muss das Territorium meines Berufsstandes schützen: Deshalb schaue ich eher in Richtung Gesellschaft und weniger in Richtung Natur. Aber das ist womöglich auch ein wenig subjektive Voreingenommenheit.

## Zur Strafbarkeit von Steuerhinterziehungen

*Lothar Kuhlen*

Ich möchte vorab kurz etwas dazu sagen, weshalb ich die Strafbarkeit von Steuerhinterziehungen als Thema meines Beitrags zu diesem Symposium gewählt habe. Das Steuerstrafrecht ist ja eine hoch spezialisierte Materie, in der sich nur ein überschaubarer Kreis von Experten wirklich auskennt. Ich gehöre nicht zu diesem Kreis, habe aber gleichwohl die vor zwei Jahren entbrannte Diskussion um den Ankauf von Steuer-CDs aus der Schweiz zum Anlass genommen, mich mit Problemen der fiskalisch und staatsbürgerlich gleichermaßen interessanten Steuerhinterziehung intensiver zu beschäftigen. Dabei ging es mir nicht um die natürlich auch interessante, aber nicht spezifisch steuerstrafrechtliche Frage nach der Zulässigkeit solcher Ankäufe und der Verwertbarkeit so erlangter Informationen, sondern um drei kriminalpolitische Forderungen, die damals erhoben wurden und auf die ich auch heute eingehen will.

Erstens sprachen sich führende Rechtspolitiker dafür aus, Hinterziehungen großen Ausmaßes vom Vergehen zum Verbrechen aufzustufen.

Zweitens forderte man eine Abschaffung der strafbefreienden Selbstanzeige.

Und drittens erwog man angesichts der großen Zahl damals entdeckter Hinterzieher eine Steueramnestie.

Diese Forderungen berühren Grundfragen der Steuerhinterziehung und zu eben diesen wollte ich einen etwas längeren Artikel schreiben. Das hat leider nicht funktioniert. Der Artikel wuchs sich vielmehr zu einer Monographie aus, die ich nach zwei Jahren zu Ende gebracht habe und die im Sommer 2012 erscheinen wird.<sup>1</sup> Einige Überlegungen, die dort näher ausgeführt werden, möchte ich an dieser Stelle vortragen.

1. Die Steuerhinterziehung ist heute als Vergehen ausgestaltet. Geschütztes Rechtsgut ist nach h. M. das staatliche Vermögen, was die Steuerhinterziehung mit anderen Vermögensdelikten, insbesondere dem Betrug, vergleichbar macht. Die Strafdrohung lautet auf Geldstrafe oder Freiheitsstrafe

---

<sup>1</sup> *Kuhlen*, Grundfragen der strafbaren Steuerhinterziehung, 2012.