Zeitschrift: Schweizerische Zeitschrift für Soziologie = Revue suisse de sociologie = Swiss journal of sociology

Herausgeber: Schweizerische Gesellschaft für Soziologie

Band: 31 (2005)

Heft: 3

Artikel: Soziale Beziehungen anstatt altruistisches Strafen! : Bemerkungen zu Ernst Fehrs Altruismusforschung

Autor: Leist, Anton

DOI: https://doi.org/10.5169/seals-815086

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 21.08.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Swiss Journal of Sociology, 31 (3), 2005, 567–582

Soziale Beziehungen anstatt altruistisches Strafen! Bemerkungen zu Ernst Fehrs Altruismusforschung

Anton Leist*

1 Zwei Traditionen und die Synthese von Individuellem und Sozialem

Zwei allgemeine, miteinander verbundene Fragen von Theorien sozialen Handelns beziehen sich auf den möglichen Vorrang eines der beiden Relationsglieder im Verhältnis von Handlungsmotiven und Handlungsumständen bzw. auf eine mögliche Konstanz oder mindestens Trägheit typischer menschlicher Handlungsmotive. Die soziale Welt könnte so beschaffen sein, dass konstante Handlungsmotive die Institutionen dieser Welt in höherem Mass bestimmen als sie umgekehrt von ihnen geprägt werden. Oder sie könnte so aussehen, dass sich die sozialen Institutionen in hohem Mass in die Handlungsmotive hinein abbilden und die sozialen Individuen des Alltags erst entstehen lassen. So vage und weiträumig formuliert, decken diese Alternativen den grössten Teil der ökonomischen und soziologischen Tradition ab, unter der Annahme einer relativen Konstanz der Handlungsmotive auf der ökonomischen Seite, einer eigenständigen *strukturellen* oder *systemischen* Dominanz der Institutionen auf der soziologischen.

Beide Traditionen können jedoch mit dem von ihnen gepflegten Vorrang nicht zufrieden sein, solange sich darüber die Logik des Restes der von ihnen konstruierten sozialen Welt nicht ebenfalls entschlüsseln lässt. Die individualistische Handlungstheorie muss die Logik der Institutionen erhellen, die systemische Sozialtheorie muss auch Zugang zur Logik des Einzelhandelns finden. Gemessen an der erklärten Theoriepolitik hat die ökonomische Tradition in den letzten Jahrzehnten wesentlich umfassendere Anstrengungen unternommen, um diesem Anspruch gerecht zu werden, als die soziologische. Verschiedenste Formen einer individualistisch-rationalen Sozialtheorie sind entstanden, die sich als Expansion des Homo-oeconomicus-Modells auf die soziale Welt jenseits des Markts verstehen – nicht ohne eine teilweise Revision dessen, was der Homo oeconomicus sein kann. Die Mikroökonomie und insbesondere die experimentelle Ökonomie studieren das Verhalten von Akteuren in sozialen Kontexten auf eine Weise, die sich in ihren Ergebnissen teilweise mit der Sozialpsychologie und Kleingruppensoziologie deckt. Dabei hält die individualistisch-rationale Theorie an einer Überzeugung ihrer Theorietradition jedoch fest, die sie von der soziologischen Tradition wesentlich unterscheidet und

* Anton Leist, Arbeits- und Forschungsstelle für Ethik, Ethik-Zentrum der Universität Zürich, Zollikerstrasse 117, CH–8008 Zürich, +41 1 634 85 28, E-Mail: leist@access.unizh.ch.

beide trotz zunehmender Konvergenzen voneinander trennt: der Annahme konstanter Verhaltensmotive und ihrer Bedeutung für die soziale Welt.

Der soweit geschilderte Gegensatz von individualistischer und systemischer Theorie ist allerdings zu radikal und führt zu überspannten und uneinlösbaren Reduktionsversuchen des einen Teils der sozialen Welt auf den anderen. Wenn das soziale Handeln nicht von Beginn an institutionenbildend ist, kann es im Weiteren sozial unerklärten Institutionen nur unterworfen werden; wenn Institutionen nicht von Beginn an Handlungserzeugnisse sind, sind sie in einen gemeinsamen Kern sozialer Phänomene nicht mehr rückführbar. An die individualistischen Begriffe der Handlungstheorie ist deshalb der Anspruch gestellt, nicht nur individuelles Handeln in sozialen Kontexten zu beschreiben, sondern die Institutionen dieses Kontexts (Normen, Regeln, Rollen etc.) als aus diesem Handeln, wenn auch möglicherweise indirekt, entstehend sichtbar werden zu lassen. Fruchtbare soziale Grundbegriffe sollten Bestandteile einer Theorie sein, in der Einzelhandeln und Institutionen auseinander entspringen und nicht sich gegenseitig vorgesetzt sind. Obwohl sich die ökonomische Tradition dieser Anforderung beispielsweise im Rahmen der Idee des Vertrags durchaus bewusst ist, haben viele der individualistisch-rationalen Expansionsversuche des *Homo oeconomicus* Schwierigkeiten, eine solche Syntheseanforderung zu erfüllen.

Stellvertretend soll diese Behauptung anhand der experimentellen Ergebnisse von Ernst Fehr im Folgenden belegt werden. Fehrs Untersuchungen zu *altruistischem* Verhalten suchen einerseits die traditionellen *egoistischen* Handlungsmotive im rationalen Handeln zu korrigieren, bleiben andererseits der Annahme konstanter Verhaltensmotive soweit verhaftet, dass schwer zu sehen ist, wie die als restringierend oder ermöglichend unterstellten Handlungsbedingungen ihrerseits aus demselben Handeln heraus erklärt werden sollen. Ein alternativer, weniger biologischer und normativ offener Begriff wie derjenige der *sozialen Anerkennung,* beschreibt, wie ich im Folgenden zeigen will, nicht nur die von Fehr studierten Verhaltensweisen phänomenal angemessener, er wird auch der geschilderten Doppelanforderung an Handlungsbegriffe besser gerecht. Nur wenn sich ein solches Forschungsprogramm von seinen stillschweigend biologisch-anthropologischen Verhaltensannahmen trennt, vermag es menschliches Handeln, in diesem Fall auch moralisches Handeln, zufriedenstellend zu erklären.

2 Altruismus und Anthropologie

Zur traditionsgebundenen Eigenart der experimentell-ökonomischen Forschung gehört, dass übergreifende und programmatische Annahmen und Ziele nicht explizit verteidigt oder auch nur genannt werden. In gewisser Weise ist das berechtigt, weil sich für empirisch-experimentell vorgehende Forscher ‹philosophische› Grund-

prämissen aus der Fruchtbarkeit ihrer Forschung ergeben müssen und nicht abstrakt verteidigen lassen. Übersehen wird dabei aber leicht, wie umfassend und radikal die Theorieziele eigentlich sind. Die Prägnanz dieser Ziele speziell in den Untersuchungen von Fehr und Mitarbeitern lässt sich, in Interpretation einiger zentraler Texte, anhand von zwei Thesen wiedergeben, einer Altruismusthese und einer Anthropologiethese. Mit *Altruismusthese* meine ich alle Versuche von Fehr und Mitarbeitern, soziale Beziehungen aus einem Gegensatz zum Egoismus heraus zu begreifen.[1] Mit *Anthropologiethese* seien die Versuche zusammengefasst, in diesem auf bestimmte Weise verstandenen Altruismus eine sozial kontextfreie Gesetzmässigkeit des Handelns zu entdecken.

Das Problematische der Altruismusthese scheint mir, wie ich im Folgenden belegen will, eben der Rückbezug auf den Begriff des Egoismus. Die mit dieser Terminologie verbundene Idee ist wohl, durch Rücksicht auf den Gegensatz Egoismus/Altruismus zu einem umfassenderen als nur dem egoistischen Handlungskomplex zu kommen. Mit dieser Absicht ist allerdings eine Verwirrung bezüglich der Relevanz des Egoismusbegriffs verbunden, nach deren Aufklärung sich Egoismus und Altruismus als nur begrenzt wichtige soziale Handlungsmotive erweisen. Altruistisches Handeln benennt keineswegs ein zentrales soziales Handlungsmotiv, wie es zunächst scheinen könnte.[2] Das Problematische der Anthropologiethese liegt erstens in der bereits geschilderten Unfähigkeit, die Normativität sozialer Phänomene aus dem Handeln heraus erklären zu können, und zweitens darin bestimmtes menschliches Verhalten kontextfrei und gleichsam aus *intrinsischen* Motiven heraus beschreiben zu müssen. Wie sich zeigen wird, treffen solche Beschreibungen auf die experimentellen Situationen, aus denen Fehr seine Daten bezieht, gerade nicht zu.

Das ungewöhnliche Ausmass ihrer Theorieperspektive scheint den Vertretern der experimentellen Ökonomie nicht erklärungsbedürftig. Quasi-biologisch ist aber offensichtlich die Absicht, den Altruismus in *Menschen* (Fehr und Fischbacher, 2003) oder die *menschliche* Kooperation (Fehr und Fischbacher, 2004) zu untersuchen – und nicht etwa altruistische Verhaltenstendenzen von Ökonomiestudenten in Zürich oder Moskau, also derjenigen Subjekte, die tatsächlich untersucht werden. Die soziobiologische Perspektive, in der Fehr seine Theorie entwickelt, ist offensichtlich an die anthropologische Gesamtvermutung geknüpft, dass sich sozial relevante menschliche Verhaltensweisen gesellschaftstranszendent, und damit menschlich *wesensimmanent,* herausgebildet haben. *Altruistische Anthropologie* ist also sicher ein passender, wenn auch erstaunlich antiquiert wirkender Titel für Fehrs Programm.

1 Der Kürze halber werde ich im Folgenden «und Mitarbeiter» fallen lassen, so dass sich der personelle Hinweis auf Ernst Fehr immer auf die Gruppe der vor allem am Zürcher *Institut für empirische Wirtschaftsforschung* Arbeitenden bezieht.

2 Zumindest gälte das, wenn der Begriff, im Sinn von Mitleid und Solidarität, richtig angewandt würde. Von Fehr wird er abweichend angewandt.

3 Zwei Bedeutungen von «Altruismus»

Um zwei Begriffe, die in Fehrs Analysen eine Hauptrolle spielen, zu beurteilen («starke Reziprozität», «altruistisches Strafen»), ist ein Verständnis von *Altruismus* nötig. In der ökonomischen Literatur wird «Altruismus» häufig als vage bleibender Sammelbegriff für alle Motive und Verhaltensweisen benutzt, die nicht *egoistisch* oder *selbstinteressiert* sind. Damit werden der Begriff und das benannte Verhalten meist nicht inhaltlich präzisiert, sondern nur in einem Gegensatzverhältnis abgegrenzt, wobei nicht selten offen bleibt, inwieweit sie indirekt durch diesen Gegensatz auch inhaltlich definiert sind. Durch diese unklaren begrifflichen Verhältnisse bleibt weitgehend offen, was mit altruistischem Handeln überhaupt gemeint ist, und diese Beobachtung trifft auch auf die Rede von Altruismus in den Arbeiten von Fehr zu.

Um zu wissen, worüber wir bei *altruistischem* Handeln reden, ist es nötig, eine Zweideutigkeit aufzuklären, die sich dadurch ergibt, dass die beiden Gegenbegriffe «egoistisches» und «selbstinteressiertes» Handeln nicht dasselbe bedeuten. Den Begriff «Altruismus» in diesem Punkt zu präzisieren ist deshalb wichtig, weil sich entsprechend zwei Bedeutungen ergeben, die enger oder weiter sind, so dass die an sie geknüpften Behauptungen wenige oder viele menschliche Verhaltensweisen erfassen und dementsprechend wenig oder stark repräsentativ sind. Ausserdem ist der Begriff dann auf unterschiedliche Weise mit alternativen Beschreibungen der Struktur sozialen Handelns vereinbar oder nicht vereinbar.

Egoistisch ist ein Verhalten dann, wenn es auf eine moralisch bedenkliche Weise allein dem Akteur dient. Statt *moralisch bedenklich* könnte man auch sagen *psychisch ungesund.*[3] *Selbstinteressiertes* Verhalten wird hingegen absichtlich als von dieser amoralischen Konnotation frei verstanden. Sich ernähren ist selbstinteressiert, aber nicht egoistisch; sich fortpflanzen ist selbstinteressiert und in der Regel günstig für andere, vor allem für die Kinder und Nachkommen. Entsprechend ergeben sich zwei Verständnisweisen von Altruismus. Altruistisches Verhalten kann jedes Verhalten sein, das nicht egoistisch ist, also auch selbstinteressiertes. Oder es kann jedes Verhalten sein, das nicht egoistisch und nicht selbstinteressiert ist. Sowohl im Alltagsgebrauch wie häufig auch in den Sozialwissenschaften wird Altruismus eher in diesem zweiten Sinn, also eng verstanden und stellt damit das konträre Spiegelbild zum Egoismus dar. Während man sich im egoistischen Handeln nur

3 Sozialwissenschaftler sind konstitutionell skeptisch gegenüber dem Einbezug *moralischer* oder *medizinischer* Begriffe in die Definition sozialer Einstellungen. Aber einfache Verhaltensbezüge sind nicht ausreichend, um Egoismus dingfest zu machen. Etwas *für sich selbst* tun würde nicht ausreichen, um ein Verhalten egoistisch zu nennen. Das eigene Leben in einer lebensbedrohlichen Situation zu retten ist sicher etwas, das man nur für sich selbst tut, aber es ist nicht egoistisch. Bei einer Überfülle von Gütern andere zu beschenken ist sicher etwas, das man *für andere* tut, aber es ist nicht altruistisch. Ein normatives Element muss deshalb in die Definitionen eingeführt werden.

um sich kümmert, kümmert man sich im altruistischen nur um den anderen (mit oder ohne Nachteilen für einen selbst). Der grösste Teil unserer Handlungen ist jedoch offensichtlich weder eng egoistisch, noch eng altruistisch (in diesem Sinn), sondern liegt eher in der Mitte des Selbstinteresses. Ein weiter Begriff des Altruismus, in dem ein Interesse an sich selbst und an anderen zusammenfallen, wäre also eigentlich angemessener. Wie ich später vorschlagen will, entsprechen das Interesse an sozialer Anerkennung oder ein Sozial- oder Gemeinschaftsinteresse diesem mittleren Verständnis, in dem Selbstinteresse und Altruismus verschmelzen. Diese Verbindung wird mit dem üblichen Begriff von Altruismus jedoch nicht ausgedrückt, weil er sich vom Gegensatz zum Egoismus nicht gelöst hat.[4]

Angesichts dieser Zweideutigkeit von «altruistisch», wird sichtbar, dass die Altruismusthese in den Untersuchungen von Fehr in eine Art Dilemma führt. Entweder ist die Verwendung des Begriffs «Altruismus» berechtigt, nämlich Altruismus im engen Sinn, dann ist unwahrscheinlich, dass damit eine sozial repräsentative Verhaltensweise benannt wird. Nur für andere handeln ist der Ausnahmefall. Oder gemeint ist «Altruismus» im weiten und insbesondere das Selbstinteresse einbeziehenden Sinn. Dann fällt zwar das Bedenken der mangelnden Repräsentativität weg, aber auch der von Fehr behauptete Gegensatz zur ökonomischen Standardtheorie, die man plausiblerweise ebenfalls mit Motiven des Selbstinteresses und nicht mit Egoismus wiedergeben sollte.[5] Wie ich noch zeigen werde, ist dies die bessere Lösung.

Fehr betont als eine Neuentdeckung im Rahmen der Mechanismen sozialen Handelns weniger Altruismus im allgemeinen als eine spezielle Variante, die er *starke Reziprozität* nennt. Im nächsten Abschnitt werde ich deshalb diesen Begriff im Verhältnis zu anderen klären (4). Daran anschliessend will ich verdeutlichen, in welchem Ausmass die von Fehr untersuchten Spiele offen lassen, welche Motive dem beobachteten Verhalten eigentlich zugrunde liegen, ohne dass die Funktion der Motive aber vernachlässigbar wäre. Freilich ist ein Bezug auf Motive unerlässlich,

4 «... in most of my dealings with others of a cooperative kind, questions of benevolence or altruism simply do not arise, any more than questions of self-interest do. In my social life I cannot but be involved in reciprocal relationships, in which it may certainly be conceded that the price I have to pay for self-seeking behaviour is a loss of certain kinds of relationships. But if I want to lead a certain kind of life, with relationships of trust, friendship, and cooperation with others, then my wanting their good and my wanting my good are not two independent, discriminable desires" (MacIntyre, 1967, 466). MacIntyre setzt Selbstinteresse mit Egoismus gleich, worin ich ihm nicht folgen kann. Zumindest bei einem sozialen Verständnis des *Selbst* scheint mir *Selbstinteresse* moralisch neutral.

5 Fehr setzt den Altruismus in Gegensatz sowohl zu *egoistischen* wie *selbstinteressierten* Verhaltensweisen (s. Fehr und Fischbacher, 2002, C1; Fehr und Falk, 2002, 691), betont aber häufiger den Gegensatz zum Egoismus bzw., wohl gleichbedeutend, «extremen Selbstinteresse» (Fehr und Fischbacher, 2002, C1). Teilweise ist auch die Rede von «nicht-pekuniären» (Fehr und Falk, 2002, 688) und «materiellen» Interessen (Fehr, Fischbacher und Gächter, 2002, 2), denen gegenüber die altruistischen dann «ideell» sein müssten. Diese begrifflichen Unterschiede werden allerdings nicht systematisch beachtet.

wenn man von altruistischem Verhalten sprechen will. Darüber hinaus schlage ich vor, dass die empirischen Daten Fehrs nicht Altruismus, sondern vielmehr Phänomene sozialer Anerkennung belegen (5). Weiter untermauern die Befunde nicht unbedingt anthropologisch verankerte Sozialmotive, sondern die hohe Abhängigkeit von sozialen Kontexten, der sozialen Normalität und legitimen Erwartungen. Diese Kontextabhängigkeit wirft ein kritisches Licht auf die Anthropologiethese und die Aussagekraft der in sozialer Isolation durchgeführten Experimente (6).

Insgesamt verstehen sich meine Bemerkungen als begrifflich angeleitete Beobachtungen eines Nichtempirikers, der sich statt auf eigene Forschung nur auf die verbreitete Alltagserfahrung mit sozialen Verhaltensweisen beziehen kann. Da sich aus ihr heraus allein letztlich die Akzeptanz der sozialwissenschaftlichen Forschung ergibt, sollten solche Konfrontationen für die eher technische Forschung aber durchaus erhellend sein.

4 Starke Reziprozität – Altruismus oder Anerkennung?

Fehr benennt das von ihm entdeckte Verhaltensmuster in verschiedenen Publikationen auf dreierlei Weisen, als ‹Altruismus›, *starke Reziprozität* und als *Fairness.*[6] «Starke Reziprozität» ist ein Kunstausdruck, mit dem wir im Alltag keine klare Bedeutung verbinden. Als altruistisch bezeichnen wir im Alltag, wie bereits geschildert, am ehesten ein bedingungsloses, freiwilliges und einseitiges Geben oder Schenken, während ein faires Verhalten mehr oder weniger gleichbedeutend ist mit einem unparteilichen oder an Gleichheit orientiertem Verhalten. Fair urteilt jemand, der unparteilich, *ohne Ansehen der Person* urteilt, und das heisst positiv ausgedrückt, gleich oder egalitär. Altruistisches und faires Verhalten sind danach nicht nur nicht identisch, sie lassen sich auch nicht unterordnen, weil Altruismus nichts mit Gleichheit zu tun hat und Gleichheit nichts mit freiwilligem Geben oder Schenken. Für das Verhaltensmuster erhellend sollte deshalb eher starke Reziprozität sein. Was ist das?

Fehr definiert starke Reziprozität so (Fehr und Fischbacher, 2003, 785):

> *«Strong reciprocity is a combination of altruistic rewarding, which is a predisposition to reward others for cooperative, norm-abiding behaviours, and altruistic punishment, which is a propensity to impose sanctions on others for norm violations.»*

6 S. insbesondere Fehr und Fischbacher, 2003 und Fehr und Rockenbach, 2003 für *Altruismus*, Fehr, Fischbacher und Gächter, 2002 für *starke Reziprozität* und Falk, Fehr und Fischbacher, 2003 für *Fairness.* Häufig werden diese Begriffe aber auch miteinander kombiniert.

Eine Doppeldeutigkeit dieser Definition ergibt sich aus der Verwendung von «altruistisch» im Definiens. Ist mit «altruistic rewarding» nichts anderes gemeint als das entsprechende Verhalten oder zusätzlich zum Verhalten eine bestimmte Motivation, etwa eine Bereitschaft des Helfens und Opferns? Im ersten Fall könnte das Prädikat durch die Verhaltensbeschreibung ersetzt werden, im zweiten Fall nicht, sondern interpretiert das Verhalten auf spezielle Weise. Zwei weitere, kleinere Probleme der Definition sind die folgenden: Erstens erfasst sie nicht eine von Fehr häufig berichtete Form starker Reziprozität, den Vertrauensvorschuss zu Beginn einer Kooperation.[7] Ein Vertrauensvorschuss ist kein Belohnen für kooperatives Verhalten und ohne einen Vorschuss solcher Art käme die Kooperation nicht in Gang. Zweitens ist zu beachten, dass die Existenz von Normen bereits zu unterstellen die Aussagekraft der Verhaltensweise erheblich einschränken kann und dem Anspruch zuwiderläuft, normengenerierende Verhaltensweisen zu studieren.[8] Ich komme auf diesen Punkt später (s. 6.) zurück. Am wichtigsten scheint aber insgesamt, wie mit der Zweideutigkeit von Verhalten ohne bzw. mit Motivbeschreibung bei altruistischen Handlungen zu verfahren ist.

Entsprechend der Zweideutigkeit können wir wählen zwischen zwei Definitionen starker Reziprozität, einer ohne und einer mit (als solcher genannter) altruistischer Motivation:

- SR1: Stark reziprok verhalten sich Akteure gegeneinander, wenn sie freiwillig Kosten zugunsten der, im Rahmen der oder in Reaktion auf die Kooperation mit anderen tragen.
- SR2: Stark reziprok verhalten sich Akteure gegeneinander, wenn sie moralisch (Altruismus, Fairness) motiviert sind, Kosten in und zugunsten der Kooperation zu tragen.

SR1 erfasst das Eingehen von Risiken zugunsten einer zukünftigen Kooperation, das Inkaufnehmen von Kosten im Rahmen und insbesondere am Ende einer Kooperation sowie die Kosten, die in Anschluss an die Kooperation getragen werden, wie etwa im nachträglichen Strafen des Kooperationspartners. Wie zwingend es ist, die so zusammengefassten Verhaltensweisen *reziprok* zu nennen, ist eine Frage, die man vielleicht nicht vertiefen sollte. Immerhin handeln alle erfassten Akteure irgendwie *in Bezug* auf andere, wenn sie auch nicht unbedingt erwarten, dass die anderen ihnen gegenüber handeln. So haben die nachträglich strafenden

7 So etwa das häufig wiederkehrende (angesichts von Massenarbeitslosigkeit allerdings unverständlich anachronistische) Beispiel des grosszügig zahlenden Unternehmers, der seine Arbeiter zu verstärkter Arbeitsmoral motivieren will. S. Fehr und Falk, 2002, 690; Fehr und Fischbacher, 2002. An einer Stelle definieren Fehr und Mitarbeiter starke Reziprozität konditional, so dass der Einstieg in die Kooperation *kein* Beispiel starker Reziprozität wäre: «The kindness of a strong reciprocator is thus *conditional* on the perceived kindness of the other player.» (s. Fehr, Fischbacher und Gächter 2002, 3, 4) Danach könnte der stark-reziproke Unternehmer keine Vorleistung erbringen.

8 Manchmal zielt Fehr auf diesen Anspruch: Fehr und Fischbacher, 2004.

Akteure keine Erwartung, dass ihnen gegenüber noch etwas geschieht, da sie ja ihre Kooperationsinteressen aufgegeben haben und deshalb nicht mehr verletzbar sind.

Obwohl die mit SR1 erfassten Verhaltensweisen darin gleich sind, dass sie eine Bereitschaft der Übernahme von Kosten zeigen, vereinen sie sich nur unter der zusätzlichen Annahme zu einem Verhaltenssyndrom, dass ebenfalls ein gemeinsamer oder mindestens zusammenhängender Motivkomplex hinter der Übernahme von Kosten steht. Andernfalls könnten sehr verschiedene Motive die äusserlich ähnliche Bereitschaft der Übernahme von Kosten tragen. Vor allem um seine Erklärungen gegen komplexere Versionen von Selbstinteresse abzugrenzen, legt Fehr tatsächlich eine gemeinsame Motivation hinter den Verhaltensweisen nahe. Damit ergibt sich die zweite Definition, SR2. Allerdings wirft SR2 ein Interpretationsproblem der beobachteten Verhaltensweisen auf, bei dem spezifische moralische Beschreibungen und Begriffe nicht mehr ausgeklammert werden können, sondern für den Sinn des Verhaltens im Licht von Akteurabsichten unverzichtbar sind. Um ein Beispiel für die Relevanz der Begriffe innerhalb von Akteurabsichten zu geben: Altruismus und Fairness sind, wie ich vorhin zeigte, nicht gleichbedeutend und auch nicht gegenseitig unterordbar, so dass SR2 in dieser Formulierung entweder eine sehr spezielle Handlungskonstellation anspricht oder zwei verschiedene Klassen von Motivation einbeziehen müsste. *Wie der Akteur einen Begriff versteht,* ist jedenfalls wichtig für seine Absicht, seine Absicht bzw. sein Motiv wichtig für den Charakter der Handlung.

Inwieweit berücksichtigt Fehr das Problem der Motivation und überführt damit Belege für SR1 in Belege für SR2? SR1 ist motivational offen und enthält keine klare Beschreibung altruistischen Handelns. Tatsächlich unterstützen Fehrs empirische Belege – wie ich im nächsten Abschnitt zeigen will – weitgehend *einzig* SR1 und lassen damit unerklärt, inwieweit das experimentell ermittelte Verhalten ein SR2-Verhalten ist. Die durchgängige Behauptung in diesem Forschungsprogramm, man könne altruistische Verhaltensweisen belegen, ist also ausgesprochen irreführend.

5 Welche Motive hinter welchen Handlungen?

Ein SR1-Handeln ist in dem Sinn vorteilhaft, dass es in den Folgen Kooperation verstärkt. Damit ist es aber noch nicht klar altruistisch, weil das Inkaufnehmen von Kosten sowohl mit egoistischem wie selbstinteressiertem Handeln vereinbar ist. Wenn jemand aus sadistischen Motiven andere straft, kann er sowohl selbst Kosten tragen wie die Kooperation unter speziellen Umständen sogar fördern. Diese und andere Möglichkeiten verdeutlichen, dass es auf die Motive ankommt, um von altruistischem Handeln zu sprechen. Fast alle Experimente von Fehr sind jedoch in ihren Aussagen über die Motive unklar.

Die wichtigsten Motive hinter Handlungen, die dem Akteur Kosten bereiten und die Kooperation fördern, dürften die folgenden sein: Akteure handeln

a) weil sie zukünftige Vorteile für sich erzeugen wollen – *indirekter Altruismus;*
b) weil sie soziale Anerkennung erhalten wollen – *Anerkennung;*
c) weil sie das Wohl von anderen / Gemeinwohl fördern wollen – *direkter Altruismus;*
d) weil sie von Fairness und Gleichheit überzeugt sind – *Fairness.*

Nach Fehr ist die (selbstinteressierte) Motivation (a) dadurch ausgeschlossen, dass die Teilnehmer sich nur ein einziges Mal begegnen. Damit ist nicht ausgeschlossen, dass manche Teilnehmer dennoch glauben, sie könnten sich weiter begegnen, aber vermutlich ist diese Wahrscheinlichkeit gering. Immerhin bleiben drei sehr verschiedene Arten von Motiven, unter denen nur c das Prädikat «altruistisch» verdient. Wie diese drei Arten von Motiven zusammenhängen, ist weder durch eine philosophische Analyse der Begriffe, noch allein durch Verhaltensbeobachtungen, sondern nur durch eine sozialpsychologische Theorie der moralischen Motivation zu klären. Meines Erachtens ist die Rede vom Altruismus, an der Fehr seine Untersuchungen orientiert, für das Entwickeln von klar einzuordnenden Forschungshypothesen bei einer solchen Anforderung eher ungeeignet, weil sie eben zu stark am differenzierungsfeindlichen Gegensatz von Altruismus/Egoismus orientiert ist, während die moralischen Motive schon begrifflich gesehen ein komplexeres Feld bilden als einen Gegensatz oder ein einfaches Kontinuum. Das wichtigste Element einer sozialpsychologischen Theorie dürfte die soziale Anerkennung (b) sein, auf die ich gleich noch einmal zurückkomme.

Wie belegt Fehr, dass gerade die Motive c und d wirksam sind? Werfen wir einen Blick auf die drei elementarsten Klassen der untersuchten Situationen.

Ultimatum Game:

Die Verhaltensweisen im Ultimatumspiel belegen die Bereitschaft, andere unter Kosten für sich selbst zu bestrafen, sowie die Fähigkeit, diese Bereitschaft oder Drohung hypothetisch vorwegzunehmen.[9] Sie belegen die Fähigkeit, die moralischen Reaktionen des Mitspielers zutreffend einzuschätzen und vorherzusagen. Dabei ist zumindest in Gesellschaften mit Fairnessstandards ein Fairnessbewusstsein am Werk, mit dem sich in etwa der Grenzpunkt einschätzen lässt, jenseits dessen der Mitspieler strafend reagieren wird. Das jedenfalls ist die vorherrschende Beschreibung der Vorgänge im Ultimatumspiel. Als *altruistisch* wird dabei die Bereitschaft des Respondenten bezeichnet, bei Zuteilungen unterhalb des Grenzpunkts die Auszahlung zu verhindern.

Was lässt sich aus dieser Beschreibung über die Handlungsmotive sagen? So gut wie nichts! Unklar ist, jedenfalls in der einfachsten Version, wann ein Drohspiel vorliegt oder wann es sich um eine schlicht faire Verteilung handelt. Eine schlicht

9 Varianten des Ultimatumspiels werden in fast allen Publikationen Fehrs zitiert und stellen den empirischen Hauptpfeiler dar, auf dem das Forschungsprogramm beruht.

faire Verteilung würde über Motive des Responders bestenfalls aussagen, dass er die faire Verteilung nicht als verletzend empfindet. Soweit der Responder straft, bleibt offen, welches der Motive b–d ihn dabei antreibt. Unter allen drei Motiven hat dasjenige der sozialen Anerkennung meines Erachtens die bei weitem höchste Plausibilität. Der Responder straft, weil er sich missachtet und verletzt fühlt. Warum er altruistisch strafen sollte, ist unklar. Bei einmaligen Spielen bleibt das Wohl unklar, zugunsten dessen der Responder strafen sollte. Ob die Strafe den Verteiler *moralisch leutert,* ist höchst ungewiss. Das lässt ein Inkaufnehmen des eigenen Verlusts zugunsten eines ungewissen Vorteils für andere leicht heroisch erscheinen. Ebenso unplausibel erscheint, dass der Responder straft, nur *um der Fairness genüge zu tun,* also zu strafen, weil unfaires Verteilen *Strafe verdient,* ungeachtet der Folgen. Damit fallen die Motive des Strafens *um des Wohls* bzw. *um der Fairness willen* weg. Und damit ist schwer zu sehen, was das Handeln des Responders altruistisch oder fair machen soll.

Wenn somit die Motive c und d ziemlich unwahrscheinlich sind, spricht vieles für das Motiv der sozialen Anerkennung, der Anerkennung als gleich gewichtigem Handlungspartner. Dieses Motiv lässt unmittelbar die Bereitschaft zum Strafen bei verweigerter Anerkennung verständlich werden. Das Motiv sozialer Anerkennung ist aber ein Motiv des Selbstinteresses und nicht des Altruismus. Sozial anerkannt zu werden, als gleich geachtet zu werden, ist zutiefst in unserem Interesse. Nur wenn man Selbstinteresse mit pekuniären, materiellen, egoistischen Interessen gleichsetzt, kann man das übersehen.

Öffentliche Güter:

In Öffentlichen-Gütern-Spielen wird, anders als im einmaligen Ultimatumspiel, das Verhalten der Mitspieler über wiederholte Spielrunden durch *altruistisches Strafen* in die Richtung kooperativen Verhaltens gezwungen (s. Fehr und Fischbacher, 2003, 787; Fehr und Gächter 2002). Als Beleg für altruistisches Verhalten wird gesehen, dass sich die stark reziproken Spieler selbst keine Vorteile aus der zukünftigen Kooperation ausrechnen können, sondern die kooperativen Lerneffekte anderen zugute kommen.

«The act of punishment does provide a material benefit for the future interaction partners of the punished subject but not for the punisher. Thus, 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. For this reason the act of punishment is an altruistic act.» (Fehr und Gächter, 2002, 139)

Zu den Motiven hat Fehr ermittelt, dass es sich um «starke negative Gefühle» handelt (Fehr und Gächter, 2002, 139). So unklar diese Beschreibung ist, legt sie doch eher nahe, dass die Strafenden aus einer Art Rache- oder Vergeltungsmotiv heraus handeln. Von einem altruistischen Motiv müsste man sagen, es sei *positiv,* nicht *negativ,* nämlich positiv auf das Gemeinwohl bezogen. Wiederum ist deshalb

eher naheliegend, dass die *negativen Gefühle* Gefühle persönlichen Missachtetwerdens ausdrücken, dass also das Motiv der sozialen Anerkennung eine zentrale Rolle spielt.

Könnte man in diesem Fall auf die Frage der Motive aber nicht ganz verzichten und den eindeutig belegten Umstand *altruistisch* nennen, dass den künftigen Handlungspartnern Vorteile entstehen, die von den früheren Strafenden verursacht werden, wenn auch vielleicht nicht beabsichtigt? Das Problem ist, dass auch selbstinteressiertes und sogar egoistisches Handeln häufig anderen Vorteile bringt. Die jetzige Investition eines Pharmaunternehmens in Forschung wird zukünftigen Patienten Vorteile bringen, ist aber nicht altruistisch. Das verbreitete Kaufen eines Guts kann das Gut auf dem Markt billiger werden lassen, so dass den künftigen Käufern durch das vermehrte Kaufen Vorteile entstehen. Dennoch würden wir nicht vom altruistischen Handeln der früheren Käufer sprechen. Ohne Bezug auf die Motive kommt das Prädikat «altruistisch» also nicht aus.

Vertrauensbildung:

In einer interessanten Klasse von Spielen kooperieren zwei Spieler so, dass ein *Verwalter* die Geldbeträge eines *Investors* vermehren kann und es ihm frei steht, den erzielten Gewinn anteilig an den Investor zurückzuzahlen. Dabei zeigt sich, dass das Rückzahlungsverhalten des Verwalters mit dem Ausmass korreliert, in dem der Investor Sanktionen für den Fall zu geringer Rückzahlungen plant (wobei die Sanktionen den Gewinneffekt bei unterlassener Rückzahlung nicht annullieren können): die *Rückzahlungsmoral* steigt, wenn der Investor freiwillig auf angedrohte Sanktionen verzichtet. Allerdings verzichten zwei Drittel der Investoren selbst dann nicht auf die Sanktionsmöglichkeit, wenn sie über diesen Effekt informiert werden (Fehr und Rockenbach, 2003). Dieses tendenziell selbstschädigende Verhalten der Investoren wird als Beleg für starke Reziprozität gesehen.

Das ist ein Beispiel für ziemlich komplexe soziale Beziehungen zwischen Vertrauen und Sanktionsdrohung. Allerdings scheinen die angebotenen Erklärungen (Fehr und Rockenbach, 2003, 139) widersprüchlich. Warum halten die über die Folgen ihres Handelns informierten Investoren an der nachteiligen Sanktionsmöglichkeit fest? Zu bedenken ist, dass die Erklärung irrational erscheint, sie würden ihren vorhersehbaren Verlust dem Wunsch nach fairem Strafen opfern (Fehr und Rockenbach, 2003); denn indem sie an der Sanktionsdrohung festhalten, machen sie Fairnessurteile überhaupt erst nötig. Beunruhigend ist ausserdem der Konflikt zwischen dem hier beobachteten kooperationsreduzierenden Effekt der Sanktionen und dem gleichen Effekt bei öffentlichen Gütern. Nach Fehr soll Fairness bei öffentlichen Gütern diesen Unterschied erklären (Fehr und Rockenbach, 2003, 140). Dieselbe Erklärung müsste aber auch bei den Zweierbeziehungen gelten, es sei denn, dass sie dort von den persönlichen Gefühlen ausser Kraft gesetzt wird. Dann müsste eine eher selbstinteressierte, an b orientierte Erklärung

allerdings auch für das Ultimatumspiel gelten, das ja ebenso ein Zweipersonenspiel ist. Die Gesamtmenge dieser Erklärungen scheint nicht kohärent.

Diese Beispiele belegen eine erhebliche Kluft zwischen den empirischen Daten zu Verhaltensweisen in diversen Spielsituationen und den möglichen Motiven hinter den Verhaltensweisen. Erstaunlich scheint mir, dass die Handlungsmotive, trotz des grossen Aufwands an empirischer Forschung, in den untersuchten Spielen weitgehend im Dunkeln bleiben. Dadurch ist es ziemlich schwierig, die Aussagekraft dieser Spiele für moralische Verhaltensweisen im Alltag einzuschätzen. Der behauptete Vorteil der Laborexperimente (Fehr und Fischbacher, 2003, 785) wird meines Erachtens durch die mangelnde Übertragbarkeit auf den sozialen Alltag stark eingeschränkt.

Mitverantwortlich für das erhebliche Interpretationsproblem scheint auch, dass die spezielle Verhaltensweise des Strafens in diesen Spielen eine so herausragende Rolle spielt und die Besonderheit dieser Verhaltensweise methodisch nicht reflektiert wird. Tatsächlich ist die strafende Reaktion eine implizit mehrdeutige und damit nur begrenzt aussagekräftige Art von Handlung. Sie ist mehrdeutig, weil sie nicht uneingeschränkt altruistisch sein kann. Der zu Strafende muss geschädigt werden, damit ihm oder anderen Vorteile entstehen. Die Motive beim Strafen müssen deshalb komplex sein und legen Anteile wie Rache, Vergeltung, Gerechtigkeit, aber sicher auch Selbstinteresse nahe. Soweit die empirischen Studien positive, etwa vertrauensbildende soziale Interessen erkennen lassen, tragen sie Züge von Motiven wie a und b. Damit steht der Altruismusanspruch in Fehrs Untersuchungen insgesamt auf einer sehr schmalen empirischen Basis.

Motive wie Rache, Vergeltung, Gerechtigkeit und Strafe verweisen allerdings zusätzlich auf einen sozialpsychologisch geräumigeren und damit angemesseneren Beschreibungsrahmen als der enge Altruismus es sein kann: auf den der sozialen Anerkennung. Fehr entgeht diese sozialpsychologisch wesentlich wichtigere Form von Beziehungen, weil er in vielen Experimenten, in Konflikt mit der sonst vorgebrachten Kritik am *materiellen* Selbstinteresse, das altruistische Strafen nur mit «echtem Geld» (Fehr und Fischbacher, 2003, 785) konzeptualisiert, während doch mindestens in wohlhabenden Gesellschaften nicht-materielle Sanktionen, eben grob gesagt der Entzug von Anerkennung, wichtiger sind als das Vorenthalten von Geld. Das Vorenthalten von Geld ist eine spezielle Version sanktionierenden Verhaltens, das unter sonst Unbekannten durch ein moralisches Bewusstsein ermöglicht wird. Mit *sozialer Anerkennung* ist ein Beachten des anderen bzw. ein Beachtetwerden durch den anderen als Handlungspartner gemeint. Diese Anerkennung ist elementarer als ein Sinn für Fairness oder Gleichheit, insofern sie sich als egalitäre Anerkennung manifestieren kann, aber nicht muss. Ein Sinn für den anderen um seiner selbst willen muss nicht egalitär, sondern kann auch hierarchisch sein. Alternativ zu Fehrs Terminologie könnte man deshalb von einem genuinen Sozialinteresse in Gestalt des sozialen Anerkanntwerdens sprechen.

Ein solches Interesse ist in der Sozialpsychologie und der philosophischen Handlungstheorie nicht unbekannt, und kann ausser mit dem Begriff der Anerkennung auch mit Begriffen wie *sozialer Identität, sozialem Beachten, Achtung, sozialer Wahrnehmung, sozialer Akzeptanz* und anderen beschrieben werden.[10] Gemeinsam ist diesen begrifflichen und theoretischen Versuchen, ein Bedürfnis des sozialen Ernstgenommenwerdens zu benennen, das im Unterschied zum emotiven Spiel des Gebens beim Altruismus deshalb wichtiger sowohl für soziale Personen wie für Beziehungen und Gesellschaften ist, weil mit ihm soziale Personen überhaupt erst entstehen. Personen sind wie Gesellschaften normativ strukturiert, und diese eigentliche *anthropologische* Einsicht kann die Rede vom Altruismus nicht einfangen.

6 Anthropologiethese und Normabhängigkeit

«If we randomly pick two human strangers from a modern society and give them the chance to engage in repeated anonymous exchanges in a laboratory experiment, there is a high probability that reciprocally altruistic behaviour will emerge spontaneously» (Fehr und Fischbacher, 2003, 785).

Äusserungen wie die eben zitierte legen in Übereinstimmung mit dem Vertrauen in die Methode isolierter und hoch-abstrakter Spielsituationen die Botschaft nahe, menschliches Verhalten beruhe zu einem wesentlichen Teil auf anthropologisch tief verankerten Verhaltenstendenzen. Dass es solche Tendenzen tatsächlich gibt, ist sicher unstrittig. Interessant ist vielmehr, wie speziell oder allgemein sie beschaffen sind und welche Anteile die soziale Umgebung daran hat, dass sie sich auf bestimmte Weise äussern. In Übereinstimmung mit meiner bisherigen Kritik scheint mir eher zutreffend, dass Menschen ein tiefsitzendes Sozialinteresse, also ein Bedürfnis nach sozialen Beziehungen um ihrer selbst willen besitzen, das Teil des Selbstinteresses ist und mit dem Egoismus in Konflikt steht. Diese tiefsitzende Sozialmotivation mit *Altruismus* und *Fairness* zu identifizieren, führt jedoch zu missverständlichen oder schlicht falschen Behauptungen.

Würde man behaupten, dass Menschen unabhängig von beeinflussenden sozialen Kontexten altruistisch handeln, könnte man entweder die für durchschnittliches Handeln wenig aussagekräftigen Weisen des engen Altruismus meinen (Geben und Schenken), oder irreführend das nicht-egoistische Handeln (weiter Altruismus). Tatsächlich handeln Menschen häufig *sozial,* also im Interesse auch von anderen, verstehen es aber nicht als altruistisch, sondern als sozialinteressiert. Würde man

10 In der Philosophie geht die Anerkennungsterminologie auf Mandeville, Rousseau und Hegel zurück, in der Sozialpsychologie auf G. H. Mead und den Symbolischen Interaktionismus (S. Honneth, 1992). Ob es tatsächlich eine *Ökonomie* der Anerkennung geben kann, wie Brennan und Pettit (2004) behaupten, ist eine interessante Frage. Allerdings trifft es nicht zu, dass die Anerkennung von A notwendig B entzogen werden muss, und vice versa.

behaupten, dass Menschen unabhängig von beeinflussenden sozialen Kontexten fair handeln, müsste man annehmen, dass moralische Gleichheit eine anthropologische Verhaltenstendenz und keine kulturelle Errungenschaft sei. Die ethnologischen Untersuchungen zum Ultimatumspiel in 15 Naturvölkern widerlegen diese Vermutung (Henrich et al., 2004). Egalitäre Verteilungen sind bei diesem Spiel in westlichen Gesellschaften häufiger als in nicht-westlichen. Damit aktualisiert sich die früher angesprochene Unterscheidung zwischen Spielsituationen ohne und mit vorausgesetzten Normen. Die von Fehr beobachteten Verhaltenstendenzen sind nicht einfach solche der Fairness, sondern der Reaktion auf eine Verletzung von Fairnessnormen, und können je nach Verständnis und Geltung von Normen sehr verschieden ausfallen.

Für (westliche) Gesellschaften aussagerelevant sind allerdings weniger abstrakte Situationen des ungleichen Teilens, als vielmehr Situationen der ungleichen Anerkennung in bestimmten Lebenslagen, persönlichen Identitäten und sozialen Rollen. Dass die Teilungsvorstellungen im abstrakten Ultimatumspiel annähernd egalitär sind, sagt noch wenig über die normativen Verhältnisse zwischen konkreten Menschen, etwa Männern und Frauen, weil die meisten Handlungsangebote und -reaktionen normativ entsprechend den sozialen Rollen und Verständnissen gefärbt und beeinflusst sind. Insofern dramatisiert sich die Frage, was die mit hohem Aufwand ermittelten Ergebnisse dieser Spiele für das Verständnis des alltäglichen Handelns eigentlich erbringen sollen.[11] Frauen mit egalitärem Selbstverständnis werden auf nicht-egalitäre Angebote strafend reagieren, wenn sie sich das sozial leisten können. Und Frauen mit einem nicht-egalitären Selbstverständnis werden auf dieselben Angebote zustimmend reagieren. Interessant wäre, ob die Reaktion unabhängig sein kann vom normativen Selbstverständnis und insofern eine eigenständige *anthropologische* Verhaltenstendenz anzeigt. Die eben angedeuteten Beispiele belegen allerdings eher das Gegenteil: In ihnen bestimmt das Fairnessverständnis auch die Reaktion. Und insofern bestimmt der gesellschaftliche Kontext die Verhaltensweise.

Damit deutet sich generell eine Konkurrenz von Erklärungsstrategien an. Man kann versuchen, anthropologische Dispositionen durch Abstraktion von speziellen Rahmenbedingungen zu belegen. Wenn die Spieler einander nicht in Rollen und im Rahmen persönlicher Situationen begegnen, könnten sie nur *als Menschen* aufeinander treffen und damit verhältnismässig generelle Dispositionen zeigen. Von gesellschaftlichen Determinanten des Handelns würde abstrahiert. Oder man kann versuchen, mehr oder weniger allgemeine Beziehungsformen zu unterscheiden und annehmen, dass sie die Motive und Handlungen weitgehend festlegen. Erst in einem zweiten Schritt würde man die Möglichkeiten studieren, wie die Beziehungsformen selbst durchbrochen, aktualisiert, variiert und gestaltet

11 Dabei unterstelle ich, dass ein egoistisches Verhaltensmodell zu widerlegen nicht nötig ist, nicht einmal, wie ich vermute, unter Ökonomen.

werden. Nach einer in der Soziologie üblichen Annahme findet jedes soziale Handeln auf dem Hintergrund eingespielter und normierter Beziehungsformen statt. Mir scheint, dass es nicht schwer ist, gegenüber Fehrs Vorgehensweise die Vorteile dieser zweiten Erklärungsstrategie zu erkennen.

Gegen die Abstraktionsmethode und die soziobiologische Strategie spricht die Vermutung, dass es eine die Handlungsmotive *nicht* beeinflussende Spielkonstellation kaum gibt. Auch hochabstrakte und dem Alltag gegenüber künstliche Spiele senden Botschaften an die Beteiligten und beeinflussen damit ihre Handlungen. Der für Nichtökonomen ziemlich auffällige Umstand, dass im Ultimatumspiel um *echtes Geld* gespielt wird, aktualisiert einen speziellen Verhaltenskreis, in dem die hohe Aversion gegen andere als egalitäre Verteilungen nicht mehr überrascht. Im Gegensatz zu den Interpretationen von Fehr vermute ich, dass die strafende Reaktion in der Regel durch eine Kombination von Selbstinteresse und Rachegelüsten (‹negative Emotionen›) zustande kommt. Anstatt altruistische Tendenzen zu belegen, scheinen die Reaktionen eher (im Sinn der zweiten Erklärungsstrategie) das Resultat einer unsicheren und ambivalenten Sozialbeziehung zu sein. Unterstützt wird diese Vermutung durch die Beobachtungen bei misslingender Vertrauensbildung (Fehr und Rockenbach, 2003). Effekte der Unterdrückung von Altruismus durch Handlungsweisen, wie sie mit der Kommerzialisierung von Gütern und Beziehungen verbunden sind, sind seit langem bekannt (Titmuss, 1970) und deshalb nicht überraschend.

Anstatt nur die einzelnen Akteure gegeneinander *Umwelt* sein zu lassen, wie beim Studium der Wirkung weniger stark reziproker Akteure auf viele Egoisten (Fehr, Fischbacher und Gächter, 2002, 11–20), schiene deshalb ein normativer Umweltbegriff vorteilhaft, mit dem zwischen verschiedenen Handlungssituationen, Aufgaben, Zielen und Konflikten und insbesondere unterstellten Normen unterschieden wird. Nichtökonomen verblüfft nicht, dass Leute mit normativen Überzeugungen entsprechend handeln und reagieren. Interessanter wäre vielmehr zu erfahren, wie sich bestimmte Beziehungsformen genauer zusammensetzen und welche Erwartungen und Verhaltensweisen sich im Rahmen dieser Beziehungen aufbauen und manifestieren. Die Arbeiten von Fehr sind von der Unterstellung geleitet, dass verhaltenswissenschaftliche Erklärungen im typisierten Zuordnen von Motiven und Handlungen unter bestimmten Bedingungen bestehen sollten. Wenn die Rahmenbedingungen jedoch – über die konkreten Handlungen der Partner hinaus – in hohem Mass Motive und Handlungen bestimmen, sind individualistische Erklärungen dieser Art wenig informativ und sollten durch *strukturelle Erklärungen* ergänzt werden.

Strukturelle, im Unterschied zu individualistischen Erklärungen weisen allerdings zwei Vorstellungen zurück, die mit sozialen Mikrotheorien wie derjenigen Fehrs eng verbunden sind. Erstens die Vorstellung, dass es generelle anthropologische Verhaltensmotive gibt, die sich in wandelnden sozialen Kontexten identifizieren

lassen und ein erkennbares Verhaltensprofil ausmachen. Gegen diese Annahme verweist die strukturelle Erklärung auf die Wirkung der Beziehungsformen. Mit Blick auf die Untersuchungen von Fehr lässt sich, so scheint mir, mindestens festhalten, dass die individualistische Erklärung der strukturellen nicht überlegen ist. Zweitens, und wichtiger, korrigiert eine Erklärung über Beziehungsformen eine durchgängige Blindheit individualistischer Erklärungen, wonach moralisches Verhalten notwendig auf Motive des *für mich* oder des *für andere* zurückgeführt werden muss, anstatt durch die Beziehungsform erklärt zu werden. Das Ultimatumspiel zum Paradigma sozialer Beziehungen zu erheben, trägt selbst bereits Züge eines solchen Missverständnisses sozialer Beziehungen. Ein für unser Leben wichtiger Teil unserer Beziehungen ist nicht von der Art, dass in ihnen geteilt werden muss. Eine Theorie, die ihre empirische Basis im Teilen hat, steht deshalb in Gefahr, die soziale und moralische Welt auf nicht unerhebliche Weise zu verkennen.

7 Literaturverzeichnis

Brennan, Geoffrey und Philip Pettit (2004), *The Economy of Esteem,* Oxford: OUP.

Falk, Armin; Ernst Fehr und Urs Fischbacher (2003), On the Nature of Fair Behavior, in: *Economic Inquiry,* 1, 41, 20–26.

Fehr, Ernst und Armin Falk (2002), Psychological Foundations of Incentives, in: *European Economics Review,* 4, 46, 687–724.

Fehr, Ernst und Urs Fischbacher (2002), Why Social Preferences Matter. The Impact of Non-Selfish Motives on Competition, Cooperation and Incentives, in: *Economic Journal,* 476, 112, C1–C33.

Fehr, Ernst und Urs Fischbacher (2003), The Nature of Human Altruism, in: *Nature,* 425, 785–791.

Fehr, Ernst und Urs Fischbacher (2004), Social Norms and Human Cooperation, in: *Trends in Cognitive Science,* 4, 8, 185–190.

Fehr, Ernst; Urs Fischbacher und S.Gächter (2002), Strong Reciprocity, Human Cooperation, and the Enforcement of Social Norms, in: *Human Nature,* 1, 13, 1–25.

Fehr, Ernst und Simon Gächter (2002), Altruistic Punishment in Humans, in: *Nature,* 415, 137–140.

Fehr, Ernst und Bettina Rockenbach (2003), Detrimental Effects of Sanctions on Human Altruism, in: *Nature,* 422, 137–140.

Henrich, Joseph et al., eds. (2004), *Foundations of Human Sociality. Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies,* Oxford: OUP.

Honneth, Axel (1992), *Der Kampf um Anerkennung,* Frankfurt: Suhrkamp.

MacIntyre, Alasdair (1967), Egoism and Altruism, in: Paul Edwards, ed., *Encyclopedia of Philosophy,* 2, 462–466.

Titmuss, Richard M. (1970), *The Gift Relationship,* Harmondsworth: Penguin.