

Verstehen, Verständigung, Vertrag – Ökonomik als Geistes-, Natur- und Staatswissenschaft

1. Einführung und Überblick

Die Wirtschaftswissenschaften beziehungsweise deren entscheidungslogisches Modell des menschlichen Verhaltens gelten vielen heutigen Ökonomen immer noch als eine mögliche Grundlage der Sozialwissenschaften insgesamt. Der Homo oeconomicus erscheint ihnen als ein universelles Verhaltensmodell (vgl. Frey Frey, B. S. (1990), McKenzie und Tullock, McKenzie, R. B. and G. Tullock (1984), Meckling, W. (1976)). Denkt man daran, welche Strukturen der „sozialtheoretische Röntgenapparat“ der modernen Mikroökonomik und Spieltheorie sichtbar machen und damit zu unserem Verständnis der sozialen und persönlichen Entscheidungsumwelt rationaler Individuen beitragen kann, so scheint der Wert einer ökonomischen Betrachtung sozialer Zusammenhänge intuitiv klar gegeben (paradigmatische nicht-formalwissenschaftlich orientierte Beiträge sind etwa Akerlof, G. A. (1984), Milgrom, P. and J. Roberts (1993), Schelling, T. C. (1978), Weise, P., W. Brandes, T. Eger and M. Kraft (2004)). Vor dem Hintergrund moderner Methodologie muss die Vorstellung einer Erklärung durch Rekurs auf die „Logik der Entscheidung“ gleichwohl als suspekt erscheinen. Deshalb haben Gegner einer entscheidungslogischen Orientierung immer wieder betont, dass die Wirtschaftswissenschaften auf die Sozial- und Individual-Psychologie zurückgeführt werden müssen, um eine im engeren Sinne sozialwissenschaftliche Grundlegung in

Verhaltensgesetzmäßigkeiten zu erhalten (vgl. dazu Albert, H. (1967/1998)). Es geht immer um menschliches Verhalten, das in allen Kontexten unter den gleichen Verhaltensgesetzen steht. Die eigentliche Basiswissenschaft ist demnach die Psychologie. Die Ökonomik wendet die psychologischen Gesetzmäßigkeiten nur unter speziellen Randbedingungen an.¹

Noch einen Schritt weiter gehen jene Ökonomen, die den Menschen als eine Spezies des Tierreiches sehen wollen (vgl. für einen interessanten inhaltlichen Versuch eines Brückenschlages Ofek, H. (2001)). Insbesondere die moderne Soziobiologie und die auf spieltheoretische Modellierungen gegründete Evolutionstheorie haben es ihnen schon wegen der formalen Analogien zu den gewohnten wirtschaftswissenschaftlichen Optimierungskalkülen angetan. Wieder andere meinen, dass ungeachtet der Tatsache, dass manche Tierspezies in der Lage sind, Traditionen zu bilden und diese weiterzugeben, menschliche Institutionen und insbesondere die Fähigkeit des Menschen zur Staatenbildung so fundamentale Veränderungen der Verhaltensumwelt darstellen, dass das natürliche Verhaltensinventar völlig sekundär gegenüber den „künstlichen“ sozialen Tatsachen wird (zum Gegensatz von „künstlich“ und „natürlich“ in der klassischen Tradition vgl. Heinemann, F. (1987/1945) und in der Neuzeit vor allem die Britischen Moralisten vgl. dazu z.B. Mackie, J. L. (1980), Raphael, D.-D. (1969)). Die Ökonomik müsse wesentlich Staatswissenschaft und wenn nicht das, dann doch genuin institutionell sein. Die deutsche Ordnungstheorie (vgl. Stützel, W. (1981) und speziell zum „Denken in Ordnungen“, Kliemt, H. (1991))

¹ Das Bestehen auf universeller Anwendung des gleichen Modells in allen Bereichen kennzeichnet natürlich auch die sogenannte „Neue Politische Ökonomie“ im Sinne etwa Buchanans, BUCHANAN, J. M. (1999 ff.): *The Collected Works of James M. Buchanan*. Indianapolis: Liberty Fund., doch ist hier der Universalismus gerade nicht empirisch-psychologisch, sondern entscheidungslogisch begründet.

gehört in diesen Zusammenhang ebenso wie die institutionelle Ökonomik und die ökonomische Theorie des Rechtes.

Auch heute noch und gerade auch unter einigen der zuvor genannten Theoretiker gibt es Anhänger einer im weiteren Sinne geisteswissenschaftlichen Ausrichtung der Ökonomik. Darunter finden sich Theoretiker, die einfach nur Mathematikverächter sind und für ihre persönliche Abneigung eine Rationalisierung suchen. Aber auch gerade jene, für die Ökonomik nur ein willkommener Anlass ist, Mathematik zu betreiben, berufen sich mit einigem Recht auf die klassische Spiel- und Entscheidungstheorie und deren Orientierung an der Binnenperspektive des rational handelnden Strategen. Durch diese Orientierung liegt der Bezug auf die überkommene geisteswissenschaftliche Kategorie des Verstehens und einer „verstehenden Sozialtheorie“ auf der Hand. Der Übergang von der apriorischen Entscheidungslogik zur aposteriorischen Erklärung bleibt jedoch auch ungeachtet neuerer Bestrebungen wie der sogenannten behavioral game theory, die insoweit allenfalls als Teilerfolg gelten darf, nach wie vor klärungsbedürftig (vgl. Camerer, C. (2003)).

Obschon es nach alledem an festen Überzeugungen methodologischer Art in der Ökonomik kaum mangelt und Ökonomen selten zögern, diese Überzeugungen auch zu äußern, lässt die beschriebene sehr unklare Gemengelage keine eindeutigen Vorschriften für angemessenes Vorgehen in der Ökonomik zu. Der Ökonom sollte für verschiedenste Ansätze offen sein. Unser eigener Beitrag versucht jedenfalls aus gegebenem Anlass einem gewissen Eklektizismus das Wort zu reden. Indem wir verschiedenste Aspekte von Verstehen, Verständigung und Verträgen ansprechen, hoffen

wir an konkreten Beispielen aufzuzeigen, wie eine Vielfalt von Betrachtungsweisen am Ende – ohne dass es eine feste Methode gäbe – einen einheitlichen „ökonomischen“ Zugang zu Problemen erkennen lässt. Wir glauben, dass dies die richtige Methode ist, Peter Weise zu würdigen, da wir in seinen breit gefächerten Schriften genau diese Vielfalt von Betrachtungsweisen zu erkennen glauben.

Konkret beginnen wir mit einer Diskussion von ganz allgemeinen Fragen des Verstehens in der Entscheidungs- und Spieltheorie (2.). Im nächsten Schritt wenden wir uns Signalen in der Tierwelt zu (3.). Die teilweise überraschend starken Parallelen zwischen biologischen und ökonomischen Formen der Theoriebildung im Bereich der Modellierung und Erfassung von Signalprozessen werden dann verdeutlicht (4.). Es folgen Schlussüberlegungen (5.).

2. Verstehen in der Entscheidungs- und Spieltheorie

Die Soziologie und hier vor allem die so genannte „verstehende“ Variante gilt vielen heutigen Ökonomen als Ausbund von Unwissenschaftlichkeit in der Sozialtheorie. Dieser schlechte Ruf ist nicht ganz unverdient. Ohne Zweifel trifft es zu, dass sich gerade auch Gegner des modernen wissenschaftlichen Weltbildes und der modernen wissenschaftlichen Methoden unter der Flagge des Verstehens und der so genannten „Hermeneutik“ versammelt haben. Man sollte darüber jedoch nicht verkennen, dass gerade einige der mathematisch-logisch fortgeschrittensten Bereiche sozialtheoretischer Modellbildung in einer näheren Beziehung zu den klassischen Konzepten wechselseitigen Verstehens stehen, als

gemeinhin gerade auch den Anhängern insbesondere der mathematischen Spiel- und Entscheidungstheorie bewusst ist.

Wer die „Logik menschlichen Handelns und Entscheidens“ untersuchen will, muss nicht so weit gehen wie Ludwig von Mises (vgl. Mises, L. v. (1940/80)) und annehmen, dass es aufgrund menschlicher Intentionalität eine Art von Wissen a priori über menschliches Handeln gibt, woraus wir ungeachtet von dessen apriorischem Charakter gehaltvolle Schlüsse über die soziale Realität ziehen können. Eine bescheidenere entscheidungslogische Perspektive einzunehmen, reicht völlig aus, um uns mit Fragen und Sichtweisen zu konfrontieren, die traditionellen Auffassungen von Prozessen wechselseitigen Verstehens zumindest sehr nahe stehen. Man erinnere sich nur daran, in welchem Maße die moderne Spieltheorie – jedenfalls in ihrer klassischen, durch John von Neumann und Oskar Morgenstern geprägten Variante – von dem Konzept gemeinsamen Wissens bestimmt ist. Das Wissen über das Wissen der anderen, das Ziehen von Schlüssen aus diesem Wissen, bei dem die Schließenden ihrerseits wiederum wissen, dass auch die anderen an einer Interaktion beteiligten Akteure ihre Schlüsse ziehen, steht im Mittelpunkt der Überlegungen.

Der „klassische“ Spieltheoretiker behandelt eine idealtypische Welt vollständig rationaler Wesen. Er geht davon aus, dass vollständig rationale Wesen ein zutreffendes Bild von ihrer Interaktionssituation in allen relevanten Aspekten haben und insoweit die Situation, in der sie sich

befinden, vollständig „verstehen“. Wenn sie etwas nicht wissen, dann wissen sie wenigstens das und auch was andere Interaktionspartner nicht wissen.²

Zumindest Oskar Morgenstern als einer der beiden Gründungsväter der modernen Spieltheorie war besonders daran interessiert, präzise zu verstehen, wie das Wissen seinen eigenen Gegenstand beeinflusst. Er fragte nach solchen Theorien, deren allgemeine Bekanntheit unter denen, die aufgrund der Theorie handeln können, nicht zu einer die Theorie widerlegenden Verhaltensänderung führen würde (vgl. dazu Morgenstern, O. (1972), Morgenstern, O. and G. Schwödiauer (1976), Dacey, R. (1976), Dacey, R. (1981) und ergänzend Güth, W. and H. Kliemt (2000)). Man kann behaupten, dass es Morgenstern, lange bevor Rawls diesen Begriff für ein verwandtes Phänomen einführte, darum ging, ein (interpersonales) „Überlegungsgleichgewicht“ (vgl. dazu umfassend Hahn, S. (2000)) zu bestimmen. In einem solchen Gleichgewicht hat niemand einen Grund zu einseitigen Überzeugungs- bzw. Einstellungsänderungen. Insbesondere sollte daher auch niemand einen Grund haben, seine Theorien und die aus diesen ableitbaren Verhaltensanweisungen zu ändern, wenn die Theorien allgemein bekannt und befolgt werden. Letztlich scheint Morgenstern mit seinem Fokus auf „absorbierbares theoretisches Wissen“ an einer tieferen philosophischen Begründung für das Cournot-Nash-Konzept interessiert gewesen zu sein, ohne dass ihm das selbst bewusst gewesen wäre (vgl. analog Jacobsen, H. J. (1996)).

Aus Sicht der Spieltheorie wissen auf einer obersten Modellierungsebene alle gemeinsam, was sie wissen bzw. nicht wissen, und ziehen dann daraus

² Für den klassischen Spieltheoretiker ist all dies gemeinsames Wissen unter allen Akteuren. Hinzu kommt sogar noch ein gemeinsames Wissen der Theorie des rationalen Spielens.

ihre Schlüsse (Aumann, R. J. (2000)). Ein Spiel als Gegenstand der klassischen Spieltheorie wäre überhaupt kein wohl definiertes Objekt, wenn nicht die Annahme des gemeinsamen Wissens gemacht würde. Genau in dem Sinne weiß jeder von jedem, was dieser weiß, was er über das Wissen der anderen weiß, was die anderen über das Wissen der andern wissen, dass sie es wissen und so weiter (da „unvollständige“ als unvollkommene Information modelliert werden kann, vgl. Harsanyi, J. C. (1967-8)).

Über dieses Wissen denken alle nach und zugleich sind alle im Rahmen der Annahmen der Theorie in der Lage, die Gedankengänge jedes anderen aus dessen von der gleichen Theorie bestimmten Sicht der Dinge „nachzudenken“. Alle vollkommen rationalen Individuen denken also in ihrem Nachdenken, in dem sie das Nachdenken der anderen vollkommen rationalen Individuen „nachdenken“ über das *gleiche* nach. In dem Sinne, dass er Zugang zum gleichen Gedanken hat, *versteht* in der klassischen Spieltheorie jeder jeden anderen vollkommen.

Das ist im wesentlichen auch der Begriff des Verstehens, der Theorien des Sprachverständnisses wie insbesondere den in David Lewis' modernem Klassiker „Konventionen“ (vgl. Lewis, D. (1975)) entwickelten, zugrunde liegt. Damit ist in der modernen Spiel- und Entscheidungstheorie der Brückenschlag zu im engeren Sinne geisteswissenschaftlichen Fragestellungen wechselseitigen Verständnisses gegeben. Doch nun scheint der Bezug zur Sozialwissenschaft fragwürdig. Für die meisten heutigen Spieltheoretiker ergibt sich die Einbindung in die Sozialwissenschaft über das Konzept der Präferenzen und das Wissen um diese. Die Präferenzen der Akteure sind gemeinsames Wissen, oder, wenn sie nicht allgemein bekannt

sind, ist wenigstens das Ausmaß der Unkenntnis allgemein bekannt. Unklar ist allerdings, wie das Konzept der Präferenzen selbst zu deuten ist. Die Nutzenfunktionen sind gerade so bestimmt worden, dass sie die Ordnung unter allen Alternativen nach Einbeziehung aller relevanten Gesichtspunkte einschließlich der Risikoneigungen der betroffenen Akteure repräsentieren.

Die individuelle Präferenzen repräsentierenden Nutzenfunktionen lassen jedoch gänzlich unterschiedliche Interpretationen zu. In der behavioralen Interpretation repräsentieren die Nutzenfunktionen nicht die Wünsche derjenigen, die an einer sozialen Interaktion beteiligt sind, sondern deren Handlungen (für eine nachdrückliche Kritik dieser Konzeption vgl. Sen, A. K. (1973/1982)). In der deliberativen Deutung hingegen sind Präferenzordnungen Wunsch- oder Bewertungsordnungen, die in eine Theorie des vollkommen rationalen Nachdenkens über strategische Interaktionen eingehen.

Viele Ökonomen neigen dazu, sich nicht lange bei der Frage aufzuhalten, was denn nun genau von den Nutzenfunktionen, mit denen sie beginnen, repräsentiert wird. Dennoch sind Wünsche und Bewertungen etwas anderes als die darauf aufbauenden Handlungen. Je nachdem, wie die an einer Interaktion beteiligten Individuen zufolge der ökonomischen Theorie die einem anderen oder ihnen selbst zugeordneten Nutzenfunktionen deuten, kommt es zu unterschiedlichen Perspektiven und Problemen.

Der Beginn mit „gegebenen Nutzenfunktionen“ verschleiert nur, dass das Problem eines Brückenschlages zwischen diesen zugrunde liegenden unterschiedlichen Perspektiven nach wie vor geleistet werden muss. Eine auf

Wissen und Nachdenken über gemeinsames Wissen abstellende Theorie, die sich im Anschluss an die deliberative Variante der Präferenzdeutung ergibt, ist sehr weit entfernt von den wirklichen kognitiven Fähigkeiten menschlicher Individuen (besonders deutlich zu sehen anhand von Werken wie Fagin, R., J. Y. Halpern, Y. Moses and M. Y. Vardi (1995)). Der Bezug zu einer verhaltenswissenschaftlichen Orientierung ergibt sich dann, wenn man die empirische Hypothese hinzufügt, dass die Menschen sich im allgemeinen ihren Wünschen und ihren Situationsanalysen entsprechend verhalten. Doch solange man an der Voraussetzung von unbeschränkten kognitiven und emotionalen Fähigkeiten festhält, hat man es allenfalls mit „als ob“ Theorien zu tun, deren explanativ wirksame nomologische Hypothesen anderer Art sind, als es die ergänzende empirische Brücken-Hypothese (vgl. dazu Albert, H. (1968)), wonach sich die Menschen ihren Wünschen und Situationsanalysen gemäß verhalten, suggeriert. Das letztere mag zwar plausibel sein, doch die Annahme, komplexe ökonomische Situationsanalysen könnten den Individuen als Gehalte von realen kognitiven Prozesse zugeschrieben werden, bleibt außer in Ausnahmefällen absurd.

Absurde Zuschreibungen lassen sich vermeiden, wenn man die so genannte Teilnehmerperspektive der klassischen Spieltheorie zu Gunsten einer nicht-klassischen behavioralen, beziehungsweise evolutionären Deutung aufgibt und die Interaktion von einem so genannten objektiven Standpunkt aus betrachtet (vgl. ursprünglich zum Teilnehmer- und zum objektiven Standpunkt, Strawson, P. F. (1962)). Man versucht als Theoretiker die Interaktion nicht mehr aus der Perspektive der verschiedenen Teilnehmer zu verstehen, sondern nimmt einen externen, objektiven Standpunkt zur

Interaktion ein. In gewisser Weise ist dies der Übergang von einer im weiteren Sinne geisteswissenschaftlichen zu einer im weiteren Sinne biologischen beziehungsweise verhaltenswissenschaftlichen Betrachtung. Nun geht es nicht mehr um eine Explikation „hermeneutischen“ Verstehens im Sinne gemeinsamen Wissens und des Bewusstseins davon in einem mathematisch präzisen Rahmen, sondern um Verstehen im Sinne der Kenntnis von Erklärungen, die auf der Basis nomologischer Hypothesen gebildet werden (vgl. zum Konzept der Explikation Siegart, G. (1997)).

Wenn wir davon ausgehen, dass die Ökonomik, so wie sie heute betrieben wird, „zwischen“ geisteswissenschaftlicher und biologischer Perspektive angesiedelt ist, dann liegt es nahe, dass wir uns nach der vorangehenden Betrachtung des geisteswissenschaftlichen Extrems nun zunächst dem anderen Extrem einer im weiteren Sinne verhaltens- bzw. evolutionsbiologischen Perspektive zuwenden. Über wechselseitiges Verstehen findet man in der Biologie praktisch nichts, dafür aber einiges über Signale und die Evolution von Signalen. Das letztere erscheint als angemessene biologische Betrachtung des Prozesses der Verständigung auch unter Menschen. Einige Grundgedanken der Theorie von Amotz Zahavi scheinen besonders geeignet dazu, dies näher zu „verstehen“ (vgl. Zahavi, A. and A. Zahavi (1998)).

3. Verständigung in der Tierwelt

Seit vielen Jahren hat sich Amotz Zahavi mit der Evolution von Signalen der Verständigung befasst und im Zuge dieser Überlegungen das mittlerweile

bekanntes „Handicap-Prinzip“ formuliert. Eine formal präzise Formulierung des Handicap-Ansatzes, den Zahavi bereits im Jahre 1975 vorschlug (vgl. Zahavi, A. (1975)), ist relativ kompliziert. Aber auch ohne diese später vor allem von Alan Grafen gelieferte (vgl. Grafen, A. (1990 a, (1990 b)), Siller, S. (1998), auch Gintis, H. (2000), Kap. 13, insbes. 13.6) formal präzise Darstellung scheint der Grundgedanke einleuchtend und zum intuitiven Einstieg in die biologische Sicht auf Signale besonders geeignet.

Die Grundhypothese, dass Signale der Verständigung den gleichen evolutionären Kräften unterliegen, wie alle anderen biologischen Phänomene, wird von Zahavi mit bemerkenswerter Konsequenz angewandt. Es muss dem Träger einer bestimmten Eigenschaft in der Fortpflanzungsmünze der Evolution nützen beziehungsweise die Ausbreitung der diese Eigenschaft tragenden Anlagen begünstigen, wenn sich die Signalanlage in einer Population ausbreiten soll. Die Signale, die von bestimmten Tierindividuen ausgesandt werden, müssen sich für die betreffenden sendenden Individuen als vorteilhaft erweisen. Zugleich muss es für die Empfänger der Signale von Vorteil sein, die Fähigkeit zu haben, solche Signale interpretieren und auf sie adäquat reagieren zu können.

Sowohl die Fähigkeit, das Signal zu senden, als auch die Fähigkeit, das Signal zu erkennen, setzen gewöhnlich voraus, dass ein gewisser „biologischer Aufwand“ getrieben wird. Diese Kosten können ohne evolutionären Nachteil nur getragen werden, wenn ihnen überkompensierende Erträge für diejenigen Individuen gegenüberstehen, die die Kosten aufgrund ihrer Anlagen auf sich nehmen.

Für unsere gegenwärtigen recht eingeschränkten Interessen reicht es aus, auf eines der Beispiele, das von Zahavi und Zahavi selbst prominent gleich zu Beginn ihres zusammenfassenden Buches eingeführt wird, einzugehen (vgl. zu solchen Beispielen auch Alcock, J. (2005)). Dieses Beispiel legt uns die einfache Frage vor, warum beispielsweise Gazellen, die einen Fressfeind wittern, nicht umstandslos und möglichst schnell die Flucht ergreifen. Wenn die Gazelle zunächst einige so genannte Prell-Sprünge vollführt, dann macht das sie noch auffälliger für den Fressfeind und kostet Zeit und Energie, die augenscheinlich besser auf eine direkte Durchführung der Flucht verwandt werden sollten. Im Rahmen herkömmlicher evolutionärer Erklärungsansätze ist das betreffende Verhalten vollkommen rätselhaft. Eigentlich sollte man erwarten, dass Gazellen, die sofort das „Hasen-Panier“ ergreifen, auf Dauer besser abschneiden und mehr Nachkommen haben würden. Die Anlage, keine dekorativen Sprünge zu vollführen, sollte sich daher im Anlagenpool durchsetzen. Dennoch beobachten wir die Prellsprünge und müssen daher im Rahmen eines darwinistischen Erklärungsprogramms dafür eine Erklärung abgeben. Die Frage, warum das, was wir beobachten, der Fall ist, führt uns zu der Suche nach den versteckten Vorteilen der betreffenden Verhaltensweisen.

Nach Auffassung von Zahavi und Zahavi zeigt die Gazelle durch ihr Verhalten ihrem Fressfeind auf, dass sie es sich erlauben kann, vor der Flucht noch „große Sprünge zu machen“. Sie kann sich das Handicap der Sprunganlage leisten und signalisiert damit dem Fressfeind ihre Fitness. Ein Fressfeind, der dieses Signal versteht, erkennt, dass es einen großen Aufwand an Energie bedeuten würde, die betreffende Gazelle zu jagen und dass das Risiko, bei der Jagd leer auszugehen und sehr viel Energie nutzlos

zu vergeuden, hoch ist. Der Beutejäger sollte daher grundsätzlich nach weniger gesunden Gazellen Ausschau halten. Für ihn ist es besser, eine lahme statt einer schnellen Gazelle zu jagen.

Es bleibt die Frage, warum auch die relativ benachteiligten Individuen sich an dem Signalprozess beteiligen sollten. Die Antwort besteht darin, dass es, wie auch sonst in der Evolution, auf relative Fitness ankommt und Beutejäger nach den *relativ* zu anderen leichter jagdbaren Individuen suchen. Auch eine nicht allzu fluchtfähige Gazelle sollte von einem Beutejäger in Ruhe gelassen werden, sofern eine noch langsamere und voraussichtlich besser jagdbare in der Nähe ist. Sofern also beispielsweise mehrere Gazellen aufgescheucht werden und Sprünge vollführen, wird der kluge Beutejäger aus dem Menü das Mahl auswählen, das er voraussichtlich mit größter Sicherheit und geringstem Aufwand erlangen kann.

Beutejäger, die es gelernt haben, aus solchen Signalen wie dem Sprungverhalten der von ihnen aufgeschreckten Opfertiere die richtigen Schlüsse zu ziehen, werden am Ende in der biologischen Konkurrenz mit Beutejägern ihrer eigenen Art besser abschneiden als Beutejäger, die die Signale nicht zu entschlüsseln vermögen. Daher sollte sich eine entsprechende Anlage in der Beutejägerpopulation durchsetzen. Soweit sich diese Anlage aber in der Population der Jäger durchsetzt, sollte sie sich auch verstärkt in der Population der potentiellen Opfer ausbreiten. Denn die Gazellen, die gut springen, werden weniger bejagt als jene, die weniger elegante Prellsprünge vollführen oder es gar nicht tun. Da es zu jeder Gazelle, die überhaupt noch zum Sprung in der Lage ist, stets noch eine geben kann (und in aller Regel geben wird), die schlechter oder überhaupt

nicht springen kann, und weil Beutejäger gewiss jede Gazelle intensiv bejagen werden, die gar nicht zum Sprung ansetzt, weil sie die direkte Flucht als untrügliches Zeichen der Schwäche deuten (da dieses Deutungsprogramm selbst adaptiv ist, setzt es sich durch), werden alle Gazellen, auch die schwachen, so gut sie können, springen und damit sämtlich ihren „Fitnessstyp“ offenbaren.

Im Gegensatz zu Modellierungen, die etwa Signalentwicklungen auf eine willkürliche Präferenz von Weibchen im Rahmen sexueller Selektion zurückführen („run away evolution“), hat man es hier durchaus mit nicht-willkürlichen Fitness-Signalen zu tun. Die Gazellen zeigen in erster Linie einem Feind und nicht einem potentiellen Partner ihre Fitness an und bringen den Feind damit u. U. dazu, lieber eine andere Gazelle zu jagen. Sie demonstrieren Eigenschaften, deren Besitz ganz unabhängig von der Frage der Partnerwahl vorteilhaft ist. Es sind diese von der Nachfrage der eigenen Artgenossen unabhängigen relativen Vorteile der eigenen Ausstattung, die „signalisiert“ werden; doch die Partner haben einen guten Grund, die Erkennung dieser Signale zu beherrschen, da dies die eigenen Fortpflanzungschancen fördert.

Da es uns hier allein darauf ankommt, strukturelle Analogien in Argumenten geisteswissenschaftlicher, biologischer und ökonomischer Art herauszuarbeiten, können wir es zunächst bei den intuitiven bisherigen Bemerkungen bewenden lassen und uns der Menschenwelt zuwenden.

4. Signale in der Menschenwelt

In der Menschenwelt gibt es durchaus zu den beschriebenen Jäger-Beute Signalen ähnliche Strukturen. So berichtet David Friedman (schmächtig, wenn auch nicht ganz so schmächtig wie sein Vater Milton), dass er sich in den fragwürdigeren Vierteln New Yorks früher mit einem Baseball-Schläger auf der Schulter zu bewegen pflegte. Das hätte, wie Friedman selbst annahm, einen körperlich starken Gewalttäter an sich nicht abschrecken sollen, doch musste durch Friedmans Ausrüstung mit einem Baseball-Schläger eine ältere Dame mit zwei Einkaufstaschen als Opfer deutlich an Attraktivität gewinnen.

Nach der zuvor diskutierten Konzeption hätte es womöglich auch für die ältere Dame Sinn gemacht, einen Schläger mitzunehmen in der Hoffnung, dass ein paar Damen ohne solche Instrumente unterwegs waren. Doch lohnt es sich nicht, die Analogie zu Jäger-Beute Strukturen an dieser Stelle weiter zu treiben. Denn in der Menschenwelt sind Signalprozesse vor allem in institutionellen Kontexten und damit in einem Bereich von Bedeutung, der in dieser Form unter Tieren nicht existiert. Bei den verbreiteteren Signalspielen in der Menschenwelt treten Signale in Verbindung mit spezifischen institutionellen Tatsachen vor allem vertraglicher bzw. vorvertraglicher Art auf.

In einem ersten Schritt führen wir vertragliche Signalisierungs- und Separierungsmechanismen zwar an einem konkreten Beispiel, doch letztlich in recht allgemeiner Form ein. Wir werden argumentieren, dass im konkreten Beispiel mit Kosten verbundene Signalisierungsmechanismen (also Handicaps) unter bestimmten Bedingungen evolutiv stabil sein

können. Dann wenden wir uns dem klassischen Signalisierspiel für den „Arbeitsmarkt“ (vgl. Spence, M. (1973), Damme, E. v. (1987), 289-293) als ziemlich enger Analogie zur Signalisierung in der Tierwelt zu. Wir untersuchen dieses Spiel abweichend von Spence im Anschluss an van Damme und Güth (1991) unter Verwendung von Annahmen der allgemeinen Gleichgewichtsauswahltheorie (vgl. Harsanyi, J. C. and R. Selten (1988)). Durch die Verwendung der allgemeinen Auswahltheorie können bestimmte ad hoc Annahmen von Spence aber auch der Spieltheorie wie die beliebte Hilfskonstruktion, gravierende Fehler seien weniger wahrscheinlich als weniger bedeutende (vgl. zum Beispiel Myerson, R. B. (1991), 222-230), vermieden werden.

4.1. Signalverträge im allgemeinen

Formal durchsetzbare Verträge gibt es letztlich nur unter Menschen. Sie sind, wenn man den Vertragsgedanken nicht überstrapaziert, ein institutionelles Phänomen spezifisch menschlicher sozialer Interaktion.³ Insbesondere sind vertragliche Beziehungen auf eine Weise intentional, die Beziehungen zwischen Tieren kaum aufweisen. Wie man Intentionen signalisiert, lässt sich durchaus in dem Rahmen spieltheoretischer Analysen von Signalisierungsprozessen untersuchen. Zugleich wird der institutionelle Aspekt damit deutlicher.

Eine aus der ökonomischen Vertragstheorie bekannte Konstellation, in der Verträge als solche (oder der spezifische Inhalt eines Vertrages) ähnlich wie ein Signal wirken, ist die der versteckten Information über Eigenschaften des

³ Selbst wenn de Waal von Koalitionsverträgen unter Schimpansen spricht, hat man doch die institutionelle Seite von Verträgen unter Menschen anders zu würdigen.

Informierten (des Agenten), welche der Agent dem Nichtinformierten (dem Prinzipal) durch das Angebot eines Vertrages mit bestimmten Inhalten signalisiert. Dass Inhalte von Vertragsangeboten Eigenschaften des Anbietenden signalisieren können, ist natürlich nur möglich, wenn eben diese Eigenschaften (innere Dispositionen oder äußere Möglichkeiten) es dem Agenten erleichtern, Verträge mit diesen Inhalten zu erfüllen und wenn diese Verträge dann auch insoweit durchsetzbar sind. Die Signalfunktion von Verträgen kann allerdings nur dann wirksam werden, wenn Verträge ansonsten nicht vollständig durchsetzbar sind, insbesondere wenn der Prinzipal auf einem Teil des Schadens eines eventuellen Vertragsbruchs sitzen bleibt, zum Beispiel weil Gerichte hier nur unvollständigen Schadensersatz gewähren oder weil sich die Vertragsparteien auf einen unvollständigen Schadensersatz einigen, damit der Prinzipal seinerseits Anreize hat, (nicht beobachtbaren) Aufwand zur Vermeidung der Nichterfüllung des Vertrages zu betreiben. Für den letzteren Fall wollen wir im folgenden ein einfaches Modell darstellen.

Konkret stelle man sich vor, ein Prinzipal P möchte eine große Party steigen lassen, die ihm persönlich sehr wichtig ist, während deren Misslingen ihm hohen Schaden verursacht. Da es sich um eine sehr große Party handelt, bedient er sich eines Agenten A – eines Partyservice-Anbieters. Der Agent kann nicht mit Sicherheit garantieren, dass die Party gelingen wird. Ob die Party ein Erfolg wird, hängt sowohl davon ab, wie gut der Prinzipal seinen Teil der Vorbereitungen ausführt, als auch von den Fähigkeiten des Agenten. Ist eine Party erst einmal misslungen, so lässt sich im Nachhinein nicht mehr feststellen, woran das im einzelnen lag. In unserer Analyse gehen wir von allgemein bekannter Risikoneutralität aller Beteiligten aus, was sich z.B.

durch die Annahme rechtfertigen lässt, dass alle Beteiligten mit vielen solcher Interaktionen konfrontiert werden.

Gelingt die Party, stiftet das dem P einen Nutzen von $V > 0$. Misslingt sie, erleidet er einen Schaden von $\tilde{S} > 0$. Zur Vereinfachung der Notation führen wir die Variable $S := \tilde{S} + V$ ein. Wie gut der Prinzipal zur Vorbereitung der Party beiträgt, sei durch dessen nicht beobachtbaren Aufwand e beschrieben. Die Wahrscheinlichkeit, dass die Party aus Gründen misslingt, die der Prinzipal durch hohen Vorbereitungsaufwand $e (> 0)$ hätte vermeiden können, sei durch die Funktion $q_1(e)$ gegeben, deren erste Ableitung negativ und deren zweite Ableitung positiv ist. Sind die ebenfalls nicht beobachtbaren Fähigkeiten des Agenten hoch, so sei die Wahrscheinlichkeit des Misslingens durch die Funktion $q_1(e)$ vollständig beschrieben. Sind die Fähigkeiten des Agenten niedrig, so komme noch eine Wahrscheinlichkeit von q_2 hinzu, so dass die gesamte Misslingenswahrscheinlichkeit $(q_1(e) + q_2) \in [0, 1]$ betrage.

Der Einfachheit halber unterstellen wir, dass es nur diese zwei Typen von Agenten gibt, der Populationsanteil der Agenten mit hohen Fähigkeiten sei durch $\lambda \in [0, 1]$ beschrieben. Da Agenten mit hohen Fähigkeiten („zuverlässige Agenten“) ihre Fähigkeiten auch anderweitig verwenden könnten, sind ihre (Opportunitäts-)Kosten k_z der Vorbereitung einer Party höher als die entsprechenden (Opportunitäts-)Kosten k_u der Agenten mit geringen Fähigkeiten (der „unzuverlässigen Agenten“). Wir unterstellen, dass $k_z - k_u < S q_2$ gilt, so dass der Prinzipal bereit wäre, die höheren Kosten

eines zuverlässigen Agenten zu tragen, wenn er die beiden Agententypen denn unterscheiden könnte.

Es sei p der Preis, den der Prinzipal an den Agenten für die Durchführung (streng genommen für den Versuch der Durchführung) zahlt, und T der Schadensersatz, den der Agent dem Prinzipal im Falle des Scheiterns der Party zahlt. Zudem sei zunächst unterstellt, dass es nur zuverlässige Agenten gebe. Dann ist der erwartete Gewinn des Prinzipals aus dem Vertrag gegeben durch

$$E(\pi_p) = (1 - q_1(e))V - p - q_1(e)(\tilde{S} - T) - e = V - p - q_1(e)(S - T) - e$$

und der erwartete Gewinn des zuverlässigen Agenten durch

$$E(\pi_{Az}) = p - q_1(e)T - k_z.$$

Unabhängig davon, wie der gemeinsame Gewinn aus dem Vertrag aufgeteilt wird, wäre es dann sinnvoll, diesen gemeinsamen Gewinn durch eine geschickte Vereinbarung über T und p zu maximieren. Dabei ist zu beachten, dass der nicht beobachtbare Aufwand e des Prinzipals von der Wahl von T beeinflusst wird. Da die Summe der beiden erwarteten Gewinne den Preis nicht mehr beinhaltet, ist das Maximierungsproblem also gegeben durch:

$$\max_T (V - q_1(\hat{e})S - \hat{e} - k_z) \text{ u.d.N. } \hat{e} = \arg \max_e (V - q_1(e)(S - T) - p - e)$$

Aus der Nebenbedingung folgt, dass entweder $\hat{e} = 0$ (nämlich für $(T-S)q_1'(0) \leq 1$, also insbesondere auch für $T > S$) oder $q_1'(\hat{e}) = \frac{1}{T-S}$ (sonst) gilt. Damit kann \hat{e} als Funktion $\hat{e}(\cdot)$ von T geschrieben werden, die stetig fällt, bis der Aufwand \hat{e} auf Null gesunken ist. Die erste Ableitung des Maximanden des Maximierungsproblems ist somit für $T > S + 1/q_1'(0)$ gleich Null und für $T < S + 1/q_1'(0) < S$ gegeben durch:

$$\left[-q_1'(\hat{e}(T))S - 1 \right] \frac{d\hat{e}}{dT} = \left[\frac{S}{S-T} - 1 \right] \frac{d\hat{e}}{dT} < 0.$$

Da somit der Maximand in T stetig fällt, bis $\hat{e}(T) = 0$ gilt, und ab diesem Wert von T konstant ist, folgt, dass in einer Welt ausschließlich zuverlässiger Agenten der vereinbarte Schadensersatz möglichst klein gehalten werden sollte. Um Missverständnisse zu vermeiden, werden wir im folgenden nicht mehr vom „vereinbarten Schadensersatz“ sprechen, sondern diesen als „Vertragsstrafe“ bezeichnen. Da eine immer kleinere Vertragsstrafe zu immer größerem Aufwand $\hat{e}(T)$ führt, impliziert das, dass die optimale Vertragsstrafe gleich Null ist. Das ist auch nicht weiter verwunderlich, da beide Beteiligte risiko-neutral sind und der Prinzipal der least-cost avoider ist (zu diesem Konzept vgl. z.B. Schäfer, H.-B. and C. Ott (2005), 226–228, 406–408 und Calabresi, G. (1972), 136–138).

Geben wir nun die Annahme auf, dass alle Agenten zuverlässig sind, und beachten, dass der Prinzipal den Typ des Agenten nicht erkennen kann, wohl aber die relative Häufigkeit λ der zuverlässigen Agenten kennt, so reduziert sich sein erwarteter Gewinn aus dem Abschluss eines Vertrages über die Party auf

$$E(\pi_p) = V - p - (q_1(e) + (1 - \lambda)q_2)(S - T) - e.$$

Der erwartete Gewinn des zuverlässigen Agenten bleibt

$$E(\pi_{Az}) = p - q_1(e)T - k_z$$

und der erwartete Gewinn eines unzuverlässigen Agenten ist

$$E(\pi_{Az}) = p - (q_1(e) + q_2)T - k_u.$$

Der zuverlässige Agent muss also mindestens einen Preis von $p = k_z + q_1(e)T$ fordern. Der unzuverlässige Agent könnte bei einer Vertragsstrafe von $T = 0$ diesen Preis unterbieten, würde damit aber seine Unzuverlässigkeit offenbaren. Auch der unzuverlässige Agent würde also ohne Vertragsstrafe mindestens einen Preis von $p = k_z + q_1(e)T = k_z$ fordern.

Selbst wenn der Prinzipal eine sehr starke Verhandlungsposition hat, so dass der zuverlässige Agent gerade mit einem Gewinn von Null aus der Verhandlung kommt, ist der Gewinn des Prinzipals bestenfalls

$$E(\pi_p) = V - k_z - (q_1(\hat{e}(0)) + (1 - \lambda)q_2)S - \hat{e}(0).$$

Wenn es genügend unzuverlässige Agenten gibt, d.h. wenn λ nicht zu groß ist, kann dieser Wert kleiner als Null sein. Ist das der Fall, so werden keine Verträge zwischen dem Prinzipal und einem zuverlässigen Agenten zustande kommen, denn entweder ist der Gewinn des Prinzipals negativ oder der des zuverlässigen Agenten.

Das heißt nun aber nicht, dass es unter keinen Umständen Verträge zwischen dem Prinzipal und einem zuverlässigen Agenten geben kann. Wird nämlich eine Vertragsstrafe $T > \underline{T} \equiv \frac{k_z - k_u}{q_2}$ vereinbart, gibt es ein Intervall von Preisen, die den Gewinn des unzuverlässigen Agenten unter Null drücken, ohne den Gewinn des zuverlässigen Agenten negativ werden zu lassen. Dass solche Preise existieren, ergibt sich unmittelbar, wenn wir die erwarteten Gewinne der beiden Agententypen bei einer Vertragsstrafe $T = \tilde{T} = \underline{T} + \varepsilon$ vergleichen, wobei ε ein kleiner positiver Wert ist:

$$E(\pi_{Au})|_{T=\tilde{T}} = p - (q_1(\hat{e}(\tilde{T})) + q_2)\tilde{T} - k_u = p - q_1(\hat{e}(\tilde{T}))\tilde{T} - k_z - q_2\varepsilon = E(\pi_{Az})|_{T=\tilde{T}} - q_2\varepsilon$$

Der Prinzipal wird das Vertragsangebot eines Agenten, das eine Vertragsstrafe leicht über \underline{T} und einem Preis, der nur dem zuverlässigen Agenten einen Gewinn verspricht, nur annehmen, wenn sein Gewinn, gegeben durch $V - p - q_1(\hat{e}(\underline{T}))(S - T) - \hat{e}(\underline{T})$, positiv ist, wenn der Agent also keinen zu hohen Preis verlangt. Der zuverlässige Agent kann einen ausreichend niedrigen Preis stets dann verlangen, wenn der gemeinsame Gewinn aus dem Vertrag, gegeben durch $V - k_z - q_1(\hat{e}(\underline{T}))S - \hat{e}(\underline{T})$, positiv ist.

Sind die beiden Ungleichungen

$$V - k_z - (q_1(\hat{e}(0)) + (1 - \lambda)q_2)S - \hat{e}(0) < 0 < V - k_z - q_1(\hat{e}(\underline{T}))S - \hat{e}(\underline{T})$$

erfüllt, kommt es also ausschließlich zu Vertragsschlüssen zwischen dem Prinzipal und dem zuverlässigen Agenten. Dieser kann bei einem entsprechenden Preis, der dem Prinzipal aber immer noch Gewinne beschert, eine Vertragsstrafe als ein Handicap auf sich nehmen. Dies senkt den

erwarteten Gewinn des Vertrages, doch nicht soweit, dass der Vertragsschluss sich nicht mehr lohnen würde.

Die vorangehenden Betrachtungen betreffen strategische Überlegungen, die vollständig rationale Akteure in einer Welt unterschiedlicher Typen und privater Typeninformation anstellen könnten. Alles Verhalten ist intentional und unterliegt keinerlei Anpassungsträgheit. Gehen wir im Gegensatz dazu davon aus, dass die Akteure keine Einzelfallmaximierer sind, sondern Routinen oder Regeln befolgen, können wir im engeren Sinne evolutorische Überlegungen ins Spiel bringen und damit eine noch deutlichere Brücke zu im engeren Sinne biologischen Modellen schlagen. Unterstellen wir also, dass sich das Verhalten aller Beteiligten durch Routineverhalten beschreiben lässt, welches sich von Zeit zu Zeit ändert und zwar im Erwartungswert in einer Richtung, in welcher die durchschnittlichen Gewinne des einzelnen aufgrund seiner Routineänderung steigen oder zumindest nicht sinken. Damit ist nicht ausgeschlossen, dass einzelne Routineänderungen auch in einer Richtung stattfinden, welche die durchschnittlichen Gewinne des einzelnen senkt. Die Anteile der Agententypen ändern sich aber im Laufe der Zeit dahingehend, dass der Anteil des Agententyps mit den aktuell höheren Gewinnen steigt. (Die evolutorische Dynamik ist also stochastisch, aber im Erwartungswert auszahlungsmonoton.) Durch Annahme einer unendlichen Population mit random matching kann dann eine erwartungswertmonotone deterministische Adaptionodynamik unterstellt werden.

Als Ausgangspunkt der evolutorischen Betrachtung sei ein Zustand gegeben, in dem erstens alle Prinzipale Verträge nur abschließen, wenn sie eine

Vertragsstrafe für den Fall des Ausfalls der Party enthalten, die einen Betrag $\bar{T} > \underline{T}$ überschreitet, und zweitens nur zuverlässige Agenten Verträge anbieten und zwar mit einer entsprechend hohen Vertragsstrafe. Solange die angebotenen und die verlangten Vertragsstrafen heterogen sind, werden sie sich offensichtlich in Richtung von \bar{T} senken. Die Dynamik wird aber noch weitergehen, da eine Senkung der minimal verlangten Vertragsstrafe in das Intervall (\underline{T}, \bar{T}) nicht zu niedrigeren Gewinnen führt und also erreicht werden kann. Ist der Anteil der Prinzipale, deren minimal geforderte Vertragsstrafen in diesem Intervall liegen, zu einem Zeitpunkt zufällig groß genug, dann werden auch Agenten die Höhe ihrer angebotenen Vertragsstrafe senken. Im Ergebnis werden sich angebotene und mindestens verlangte Vertragsstrafen immer mehr dem Wert \underline{T} annähern.

Solange keine unzuverlässigen Agenten auf den Markt drängen, ändert sich an dieser Aussage auch dann nichts, wenn der Wert \underline{T} unterschritten wird. Angebotene und mindestens verlangte Vertragsstrafen werden weiter sinken. Doch ist der Wert \underline{T} erst einmal unterschritten, so wird der Anteil der unzuverlässigen Agenten steigen, zum Beispiel weil solche Agenten von anderen Märkten zuwandern. Die unzuverlässigen Agenten können unter den gegebenen Bedingungen billiger anbieten als die zuverlässigen Agenten und werden sich daher ausbreiten. Ist der Anteil der unzuverlässigen Agenten groß genug, werden Prinzipale, die eine Vertragsstrafe oberhalb von \underline{T} verlangen, wieder höhere Gewinne erhalten als Prinzipale, die sich mit einer niedrigeren Vertragsstrafe zufrieden geben, so dass deren Anteil wieder steigt. Ob dadurch ein Zustand, in dem vor allem Vertragsstrafen oberhalb von \underline{T} gefordert und geboten werden und (fast) alle Agenten

zuverlässig sind, stabilisiert wird, hängt von den Einzelheiten der Dynamik ab, insbesondere von der Geschwindigkeit und Homogenität, mit der die Neigung der Prinzipale steigt, Verträge zu akzeptieren, deren Vertragsstrafe kleiner als \underline{T} ist.

Ist ein solcher Zustand stabil, stellt er ein Beispiel für die Wirksamkeit des Handicap-Prinzips im institutionellen Bereich der Menschenwelt dar: eine Population, hier die (zuverlässigen) Agenten, signalisiert durch ein Handicap, also ein Signal, welches sich hier sogar zum Nachteil *beider* Vertragsparteien auswirkt, bestimmte Eigenschaften und die Mitglieder der anderen Population, hier der Prinzipale, verstehen dieses Signal und legen es ihrem Verhalten zugrunde.

Insoweit die Prinzipale sich die Agenten danach aussuchen, welche Art von Verträgen diese anbieten, wird deutlich, dass sich zumindest in dem hier dargestellten evolutorischen Modell *screening* und *signalling* nur gemeinsam stabilisieren können. Sie sind, was das anbelangt, als zwei Seiten derselben Medaille zu betrachten – eine Einsicht, zu der in einem ganz anders strukturierten Modell auch Nöldeke und Samuelson kommen (vgl. Nöldeke, G. and L. Samuelson (1997)). -- Im Falle von Arbeitsverträgen wird die Analogie zu den Signalisierungsspielen in der Tierwelt womöglich noch augenfälliger.

4.2. Arbeitsverträge und “Fitness-Signale”

4.2.1. Das Spiel

Hier ist die einfache Situation (vgl. Spence, M. (1973)): Eine Firma F möchte einen Bewerber einstellen, den sie auf Grund ihrer a priori-Erwartung mit Wahrscheinlichkeit $\lambda \in (0,1)$ als produktiv (genormt auf Produktivität 1) einschätzt und mit der Restwahrscheinlichkeit $1-\lambda$ als unproduktiv (genormt auf Produktivität 0). Diese Einschätzung durch F sei auch beiden Typen des Bewerbers bekannt.

Vor der Einstellung kann der Bewerber einen crash course in Wirtschaftswissenschaften absolvieren. Wir bezeichnen mit $y \geq 0$ die Anzahl der Kreditpunkte, die der Bewerber im Rahmen des crash courses erwirbt. Die Absolvierung des Kurses ist kostenträchtiger oder weniger kostenträchtig für den Bewerber selbst je nachdem, ob es sich um einen „fitten“, produktiven oder weniger „fitten“, d.h. auch für die Firma weniger produktiven Besucher des crash courses handelt.

Die Bereitschaft, den Kurs auf sich zu nehmen, entspricht in etwa der Bereitschaft der Gazelle, Prellsprünge durchzuführen. So wie eine „fittige“ Gazelle geringere Kosten (Risiken) auf sich nimmt, wenn sie springt, so bedeutet es für einen produktiveren Bewerber einen geringeren Aufwand, wenn er sich dem crash course unterzieht. So wie die Prellsprünge von Gazellen pure Verschwendung sind, so ist der crash course es für jene Bewerber, die sich ihm unterziehen.

Hier enden die Analogien allerdings. Es ist anders als im Falle der Gazellen für die menschlichen Bewerber nicht mehr zwingend, dass alle tatsächlich

die Kosten auf sich nehmen. Denn die Firma ist nicht in der Lage, unterschiedliche Qualitäten direkt zu beobachten. Insbesondere kann die Firma aus dem Punktwert y nicht unmittelbar die Eignung des Kandidaten ablesen. Die Firma ist darauf angewiesen, z.B. durch geschickte Lohnfestsetzung, die unproduktiven Bewerber dazu zu bringen, möglichst keinen crash course zu belegen, weil das für diese zu teuer ist, während der Kursbesuch für die produktiven aufgrund ihrer geringeren Kosten lohnt.

Den produktiven muss es leichter fallen als den unproduktiven, das Signal zu erzeugen. Das ist so, als ob die Gazellenjäger unterhalb eines bestimmten Schwellenwertes auf jeden Fall und oberhalb eines solchen Wertes keinesfalls die Beute verfolgen.⁴ Dann würde es für Gazellen mit Qualitäten unterhalb der Schwelle auch keinen Sinn mehr machen, überhaupt Sprünge zu unternehmen, die anderen würden hingegen gerade so springen, dass sie sich von denen trennen, die keine großen Sprünge machen können.

Für den unproduktiven Typ U seien die Kosten C_U für die Anzahl y_U an Kreditpunkten einfach $C_U = y_U$, während der produktive Typ P für einen Abschluss mit y_P Kreditpunkten nur Kosten in Höhe von $C_P = \frac{y_P}{2}$ aufzubringen hat. Der Bewerber konfrontiert die Firma F mit einer take it or leave it-Lohnforderung $w \in [0,1]$, welche die Firma F nur annehmen oder ablehnen kann.

⁴ Es könnte sich hier auch um ein Intervall handeln, innerhalb dessen die Entscheidung von der Notwendigkeit eines unmittelbaren Jagderfolges abhängt.

Lehnt F ab, so verdient die Firma „0“. Der Bewerber muss seine Kosten tragen, d.h. der U-Typ verdient $-y_U$ und der P-Typ des Bewerbers $-\frac{y_P}{2}$.

Nimmt die Firma die Lohnforderung an, so verdient sie „brutto“ „1“ und „netto“ „ $1-W$ “, falls der Bewerber vom Typ P ist und „ $0-W$ “, falls sie einen Bewerber vom unproduktiven U-Typ engagiert hat. Der P-Typ des Bewerbers verdient „ $W - \frac{y_P}{2}$ “ und der U-Typ „ $W - y_U$ “.

Der Prozess, gemäß dem diese Auszahlungen festgelegt werden, ist der folgende:

- Der Zufall wählt den P-Typ bzw. den U-Typ des Bewerbers mit Wahrscheinlichkeit λ bzw. $1-\lambda$ aus, wobei $0 < \lambda < 1$. Über das Ergebnis wird ausschließlich der Bewerber selbst informiert; die Firma F kennt nur die Wahrscheinlichkeit λ .
- Der Bewerber(typ) wählt sein Ausbildungsniveau y_P bzw. y_U aus. Die Firma F erfährt die Punktzahl y des Bewerbers, aber nicht, ob es sich bei dem Wert von y um das $y_U = y$ eines U-Typs oder um das $y_P = y$ eines P-Typs handelt. Gleichzeitig bestimmt der Bewerber seine ultimative Lohnforderung W mit $0 \leq W \leq 1$.
- In Kenntnis von y und W entscheidet sich die Firma F für oder gegen den Bewerber.

Nachdem die Stufen durchlaufen wurden, kommt es zu den vorgenannten Auszahlungen. Der entscheidende Politikparameter der Firma ist die Festlegung jenes Lohnsatzes, den sie zu zahlen bereit ist. Intuitiv sollte dieser möglichst so hoch sein, dass produktive Arbeitnehmer dafür arbeiten werden, weil er deren Kosten übertrifft und so niedrig, dass die

unproduktiven von vornherein davon abgehalten werden, überhaupt einen crash course zu belegen. Betrachten wir mögliche Gleichgewichte genauer.

4.2.2. Uniform perfekte Gleichgewichte als Lösungen des Spiels

Bezeichnen $q_p(y)$ bzw. $q_U(y)$ die (a priori) Wahrscheinlichkeiten, mit denen jeweils der P- bzw. U-Typ des Bewerbers das Abschlussniveau y realisieren, dann ist die Posteriori-Produktivitätserwartung der Firma F auf der letzten Stufe des Entscheidungsprozesses für $\lambda q_p(y) + (1-\lambda)q_U(y) > 0$ eindeutig festgelegt durch

$$\mu(y) = \frac{\lambda q_p(y)}{\lambda q_p(y) + (1-\lambda)q_U(y)} 1 + \frac{(1-\lambda)q_U(y)}{\lambda q_p(y) + (1-\lambda)q_U(y)} 0 \text{ oder}$$

$$\mu(y) = \frac{\lambda q_p(y)}{\lambda q_p(y) + (1-\lambda)q_U(y)}, \text{ falls } \lambda q_p(y) + (1-\lambda)q_U(y) > 0.$$

Um sich stets auf diese Definition $\mu(y)$ beziehen zu können, betrachten Harsanyi und Selten (1988) uniform perturbierte Spiele und lösen das unperturbierte Spiel durch Grenzübergang (wenn die uniforme Perturbation gegen Null konvergiert).

Für alle $\mu(y)$ wird die optimale Lohnforderung durch $W^*(y) = \mu(y)$ bestimmt. Würden beide Typen des Bewerbers das Ausbildungsniveau y mit der gleichen Wahrscheinlichkeit $q_p(y) = q_U(y) > 0$ realisieren, ergäbe sich sogleich $W^*(y) = \lambda$. Da weder der U- noch der P-Typ freiwillig $y > 2$ wählen werden, muss also $W^*(y) = \lambda$ für $y > 2$ gelten. Wir können uns im folgenden

auf die Ableitung von $W^*(y)$ für $0 \leq y \leq 2$ beschränken. Hierfür gibt es die sogenannten pooling-Gleichgewichte mit Lohnfunktion $W^*(y)$ für alle Gewinnhöhen $\lambda - \bar{y}$ mit $\bar{y} \in [0, \lambda]$ des U-Typs wie in der folgenden Abbildung:

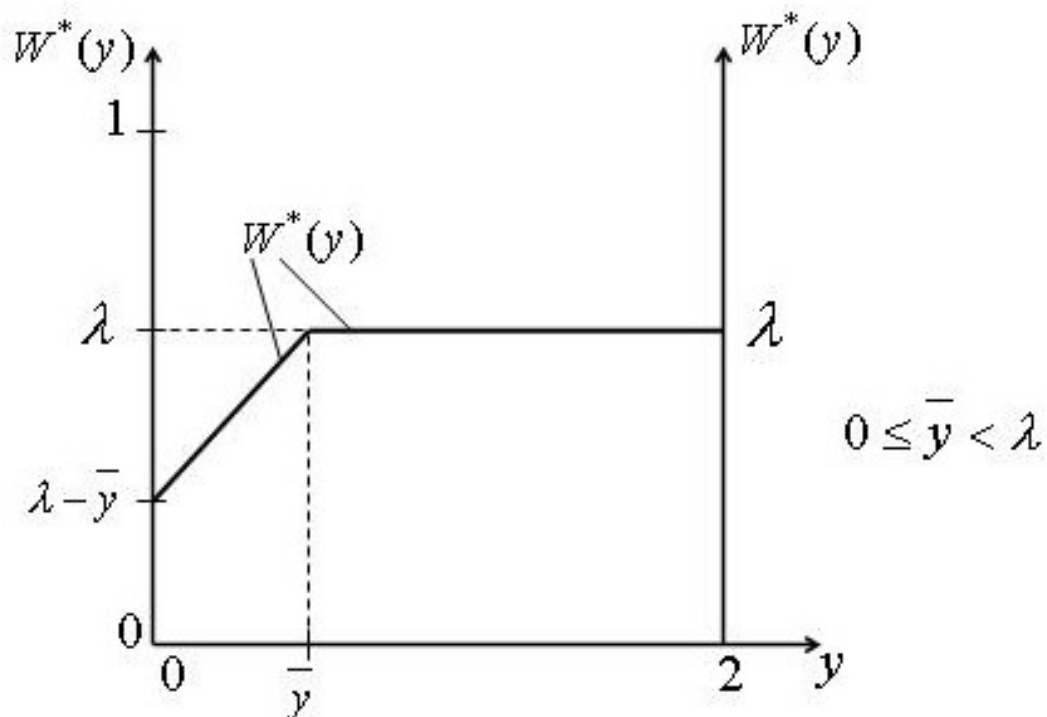


Abbildung 1

Wie durch Abbildung 1 verdeutlicht, gibt es so viele Pooling-Gleichgewichte wie Parameterwerte $\bar{y} \in [0, \lambda]$; wobei $\lambda - \bar{y}$ die Gewinnhöhe des unproduktiven U-Typs für das jeweilige Gleichgewicht bemisst. Eine solche Lohnfunktion resultiert, wenn der U-Typ Ausbildungsniveaus $y < \bar{y}$ freiwillig mit höherer Wahrscheinlichkeit $q_U(y)$ realisiert als der P-Typ, für

den $q_p(y)$ im Bereich $y < \bar{y}$ minimal ist. Offensichtlich kann durch freiwillige Wahl von $q_u(y) > \varepsilon$ die posteriori vom Arbeitgeber erwartete Arbeitsproduktivität $\mu(y)$ im Bereich $y < \bar{y}$ auf das (für die Gewinnhöhe $\lambda - \bar{y}$ des unproduktiven U-Typs) erforderliche Niveau reduziert werden, was wir noch genauer demonstrieren werden. Der P-Typ selbst wählt $y_p = \bar{y}$ mit der im uniform perturbierten Spiel maximal möglichen Wahrscheinlichkeit; d.h. $q_p(y) = \varepsilon, \forall y \neq \bar{y}$.

Wieso kann der U sich so frei im Bereich $y < \bar{y}$ verhalten? Der Grund ist, dass für den U-Typ des Bewerbers der ansteigende Bereich von $W^*(y)$ in der Abbildung mit der Indifferenzkurve des U-Typs für das Gewinnniveau $\lambda - \bar{y} \geq 0$ übereinstimmt (da $W^*(y_U) - y_U = \lambda - \bar{y} + y_U - y_U = \lambda - \bar{y}$ für alle $y_U < \bar{y}$). Dies ermöglicht es dem U-Typ, seine Wahrscheinlichkeit für $y_U < \bar{y}$ in der für $W^*(y)$ erforderlichen Weise festzulegen. Es sei $\varepsilon (> 0)$ die Minimalwahrscheinlichkeit für alle y und damit $q_p(y) = \varepsilon$ für alle $y \neq \bar{y}$ sowie $q_u(y) = \varepsilon$ für alle $y > \bar{y}$. Es folgt dann aus dem im Sinne von Abbildung 1 unterstellten Verlauf der Lohnkurve $W^*(y)$

$$\mu(y) = \frac{\lambda \varepsilon}{\lambda \varepsilon + (1 - \lambda) q_u(y)} = \lambda - \bar{y} + y$$

für alle $y < \bar{y}$ die Bedingung

$$q_u(y) = \lambda \varepsilon \frac{1 - \lambda + \bar{y} - y}{\lambda - \bar{y} + y}.$$

Offensichtlich konvergiert $q_U(y) \rightarrow 0$ für $\varepsilon \rightarrow 0$. Die Grenzlösung für das ungestörte Spiel mit $\varepsilon = 0$ ist also dadurch gekennzeichnet, dass $q_U(\bar{y}) = 1 = q_P(\bar{y})$ gilt. Wir haben damit die folgende Aussage bewiesen:

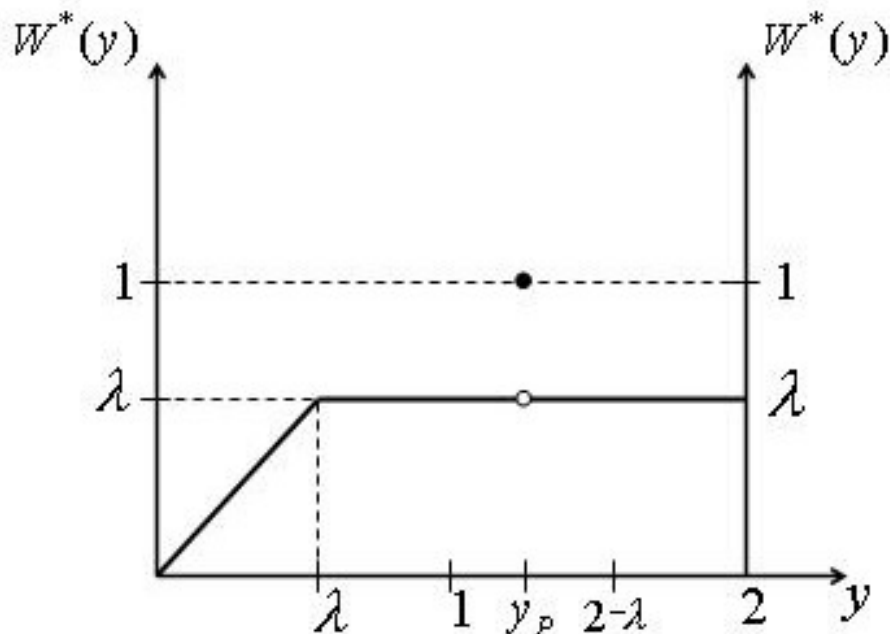
Proposition 1: Für alle \bar{y} mit $0 \leq \bar{y} < \lambda$ ist $y_U = \bar{y} = y_P$ ein uniform perfektes pooling-Gleichgewicht, in dem der U-Typ $\lambda - \bar{y}$ und der P-Typ $\lambda - \frac{\bar{y}}{2}$ verdient.

Uniform perfekt nennen wir ein solches Gleichgewicht, da es nicht durch Gleichgewichte von beliebigen Folgen perturbierter Spiele (Selten, 1975), sondern durch Folgen von Gleichgewichten uniform perturbierter Spiele als Grenzlösung des ungestörten Spiels nachgewiesen wurde. Beim Grenzübergang ändern sich zwar $q_U(y)$ und $q_P(y)$ für alle $y < \bar{y}$, aber nicht die Lohnfunktion $W^*(y)$, da sich sowohl $q_U(y)$ als auch $q_P(y)$ mit ε ändern und ihre Effekte sich bezüglich $\mu(y)$ genau aufheben (vgl. die obige Ableitung von $q_U(y)$). Das pooling-Gleichgewicht mit $\bar{y} = 0$ wird auch als bestes pooling-Gleichgewicht bezeichnet, da es minimale Ausbildungskosten aufweist.

Signalisiergleichgewichte im ε -uniform gestörten Spiel mit $\varepsilon > 0$ verlangen maximale Wahrscheinlichkeit $q_U(0)$, gegeben $q_U(y) = \varepsilon$ für $y \geq \lambda$ und $q_U(y) = \lambda \varepsilon \frac{1 - \lambda - y}{\lambda + y}$ für alle $0 < y < \lambda$. Bezüglich des U-Typs entspricht das Verhalten $\bar{y} = \lambda$ und damit einer Lohnkurve $W^*(y)$, die im Bereich $0 \leq y < \lambda$ mit der 0-Isogewinnkurve des U-Typs übereinstimmt. Um für den U-Typ jede lohnende Abweichung auf das vom P-Typ mit maximaler Wahrscheinlichkeit gewählte Ausbildungsniveau y_P auszuschließen, muss

$y_p \geq 1$ gelten, da selbst der maximale Lohnsatz von 1 Ausbildungskosten des U-Typs von $y_U = y_p \geq 1$ nicht rechtfertigen kann. Umgekehrt sollte P nicht dadurch gewinnen können, dass er $y < y_p$ statt y_p zu realisieren versucht. Da die beste solche Abweichungsmöglichkeit wegen der Kosten in Höhe von $y_p/2$ und wegen der in Abbildung 1 resultierenden Lohnkurve für $\lambda = \bar{y}$ durch $y_p = \lambda$ gegeben ist (vgl. die folgende Abbildung 2), erfordert dies

$$1 - \frac{y_p}{2} > \lambda - \frac{\lambda}{2} \text{ bzw. } y_p < 2 - \lambda.$$



Wenn nun ε gegen Null konvergiert, impliziert dies die Grenzlösung von $q_u(0)=1$ und $q_p(y_p)=1$ für ein y_p mit $1 \leq y_p < 2 - \lambda$. Wir haben die Lohnfunktion für die Grenzlösung im Sinne von $\varepsilon \rightarrow 0$ in einer analogen Abbildung zur vorherigen verdeutlicht. Abgesehen von der notwendigen

Festlegung auf $\lambda = \bar{y}$, da der U-Typ auf einen Nullgewinn reduziert ist, besteht der wesentliche Unterschied nur im Lohn $W^*(y_p) = 1$ für ein bestimmtes Ausbildungsniveau mit $1 \leq y_p < 2 - \lambda$, da nur der P-Typ dieses Ausbildungsniveau wählt und Imitation durch den U-Typ nicht zu befürchten ist, wenn dieser sich rational verhält. Es gilt daher

Proposition 2: Für alle y_p mit $1 \leq y_p < 2 - \lambda$ ist durch $q_u(0) = 1$ und $q_p(y_p) = 1$ ein uniform perfektes Signalisiergleichgewicht gegeben, in dem der U-Typ Null und der P-Typ $1 - \frac{y_p}{2}$ verdient.

Wegen der für den P-Typ des Bewerbers geringeren Ausbildungskosten wird das Gleichgewicht mit $y_p = 1$ als bestes Signalisiergleichgewicht bezeichnet. Selbst wenn man sich auf die besten Gleichgewichte beschränkt, bleibt jedoch – und das ist in unserem Zusammenhang von besonderer Bedeutung! – die Koexistenz des besten pooling- und des besten Signalisiergleichgewichtes bestehen. Wir formulieren dies als

Korollar 3: Das von uns in Anlehnung an Spence analysierte Signalisierspiel verfügt typischerweise über multiple uniform perfekte Gleichgewichte (Harsanyi und Selten, 1988), von denen zumindestens die pooling-Gleichgewichte Handicap-Phänomene leugnen.

5. Diskussion

Wenn zwei potentielle Tauschpartner einen Tausch verabreden, dann streben sie ein für sie beide gegenüber dem Status quo superiores Resultat an. Es liegt im gemeinsamen Interesse, den Austausch zu realisieren, denn

gegenüber dem Status quo würden sich beide besser stellen. Deshalb haben beide guten Grund, sich „in foro interno“ – wie Hobbes gesagt hätte (vgl. Hobbes, T. (1651/1976)) – zu wünschen, dass der Austausch zustande kommt. Jedoch würde sich jeder einzelne bei einseitiger Erfüllung des Vertrages durch den anderen noch besser stehen, als wenn beide Partner einschließlich seiner selbst erfüllen würden.

Sofern man es mit einer Einmal-Interaktion zu tun hat, ist die Einhaltung von Versprechen nicht teilspielperfekt (vgl. zu Versprechen Lahno, B. (1995)). Die Rückseite des wechselseitig vorteilhaften Tausches bildet insoweit ein Gefangenen-Dilemma-Spiel (vgl. dazu etwa auch Hardin, R. (1982), der dies bündig zusammenfasst, indem er von „prisoner’s dilemma or exchange“ spricht).

Wegen der Unsicherheit, ob Versprechen zur Zusammenarbeit auch wirklich eingehalten werden, haben menschliche Individuen guten Grund, für ihre Zusammenarbeit und ihre Austausch-Akte nach intrinsisch zur Einhaltung von Versprechen motivierten, insoweit vertrauenswürdigen Partnern zu suchen. Diese Suche kann nur dann erfolgreich sein, wenn man die vertrauenswürdigen Partner erkennen kann. Die Fähigkeit, vertrauenswürdige Partner zu erkennen, kann einerseits auf der Erkenntnisfähigkeit dessen beruhen, der einen Partner sucht, dem er vertrauen kann (vgl. dazu Güth, W., H. Kliemt and B. Peleg (1999)). Sie kann sich andererseits aus der Signalisierungs-Fähigkeit dessen herleiten, der einen Partner sucht, der ihm vertraut.

Was das Letztere anbelangt, denkt man gewöhnlich in der Literatur an die glaubwürdige Signalisierung persönlicher Eigenschaften, die sich letztlich in einem Typ beziehungsweise einer bestimmten Nutzenfunktion niederschlagen. Die Glaubwürdigkeit setzt bekanntlich voraus, dass es teuer ist, ein falsches Signal zu senden. Durch sprachliche Zeichen zu lügen, ist hingegen einfach. Daher gilt es, Bedingungen zu bestimmen, unter denen Signalfälschungen dieser Art für menschliche Individuen teuer werden.

Signalfälschungen können unter Menschen teuer werden, aufgrund von „natürlichen“ oder „künstlichen“ Mechanismen. Natürliche Mechanismen können insbesondere dadurch die Signalfälschung verteuern, dass menschliche Individuen in ihrem Entscheidungshandeln von Emotionen, die ihnen eine beliebige opportunistische Situationsdiskriminierung *nicht* gestatten, gesteuert werden (vgl. Frank, R. (1992)). Durch Nutzung der „künstlichen“ Institution des Vertrages kann man es erreichen, dass derjenige, der seine Versprechen einhalten wird, entsprechende Verträge anbietet oder annimmt. Solch eine Nutzungsmöglichkeit der künstlichen Vertragsinstitution ist allgemein bekannt, deren Nähe zu Signalprozessen in der natürlichen Tierwelt ebenfalls. Wie stark die Parallelen sind und inwieweit man sie in allgemeine Auffassungen vom Verständnis wechselseitiger Intentionen in strategischen menschlichen Interaktionssituationen mit privater Information einbetten kann, ist womöglich durch das vorangehende etwas deutlicher und konkreter illustriert worden, bedarf aber nach wie vor einer allgemeinen Untersuchung und Aufarbeitung. Die nachfolgend recht ausführlich angeführte Literatur bietet insoweit weiteres Material zur Herausarbeitung gemeinsamer struktureller Eigenschaften einer Vielzahl einschlägiger Ansätze.

6. Literatur

- AKERLOF, G. A. (1984): *An Economic Theorist's Book of Tales*. Cambridge: Cambridge University Press.
- ALBERT, H. (1967/1998): *Marktsoziologie Und Entscheidungslogik. Zur Kritik Der Reinen Ökonomik*. Tübingen: Mohr.
- (1968): *Traktat Über Kritische Vernunft*. Tübingen: Mohr.
- ALCOCK, J. (2005): *Animal Behavior. An Evolutionary Approach*. Sunderland, MA: Sinauer.
- AUMANN, R. J. (2000): *Collected Papers I & II*. Cambridge, Mass.: MIT Press.
- BUCHANAN, J. M. (1999 ff.): *The Collected Works of James M. Buchanan*. Indianapolis: Liberty Fund.
- CALABRESI, G. (1972): *The Costs of Accidents: A Legal and Economic Analysis*. New Haven: Yale University Press.
- CAMERER, C. (2003): *Behavioral Game Theory*. Princeton: Princeton University Press.
- DACEY, R. (1976): "Theory Absorption and the Testability of Economic Theory," *Zeitschrift für Nationalökonomie*, 36, 247-267.
- (1981): "Some Implications of 'Theory Absorption' for Economic Theory and the Economics of Information," in *Philosophy in Economics*, ed. by J. C. Pitt. Dordrecht: D. Reidel, 111-136.
- DAMME, E. V. (1987): *Stability and Perfection of Nash Equilibria*. Berlin / Heidelberg / New York / London / Paris / Tokyo: Springer Verlag.
- FAGIN, R., J. Y. HALPERN, Y. MOSES, and M. Y. VARDI (1995): *Reasoning About Knowledge*. Cambridge, MA / London: MIT Press.
- FRANK, R. (1992): *Die Strategische Rolle Der Emotionen*. Oldenbourg: Oldenbourg Verlag.
- FREY, B. S. (1990): *Ökonomie Ist Sozialwissenschaft. Die Anwendung Der Ökonomie Auf Neue Gebiete*. München: Vahlen.
- GINTIS, H. (2000): *Game Theory Evolving*. Princeton: Princeton University Press.
- GRAFEN, A. (1990 a): "Biological Signals as Handicaps," *Journal of Theoretical Biology*, 144, 517-546.
- (1990 b): "Sexual Selection Unhandicapped by the Fisher Process," *Journal of Theoretical Biology*, 144, 473-516.
- GÜTH, W., and H. KLIEMT (2000): "From Full to Bounded Rationality. The Limits of Unlimited Rationality," Bielefeld: Center for Interdisciplinary Research (ZiF).
- GÜTH, W., H. KLIEMT, and B. PELEG (1999): "Co-Evolution of Preferences and Information in Simple Game of Trust," *German Economic Review*, 1, 83-110.
- HAHN, S. (2000): *Überlegungsgleichgewicht(E). Prüfung Einer Rechtfertigungsmetapher*. Freiburg i.Br.: Karl Alber.
- HARDIN, R. (1982): "Exchange Theory on Strategic Basis," *Social Science Information*, 2, 251 ff.
- HARSANYI, J. C. (1967-8): "Games with Incomplete Information Played by Bayesian Players," *Management Science*, 14, 159-82, 320-34, 486-502.

- HARSANYI, J. C., and R. SELTEN (1988): *A General Theory of Equilibrium Selection in Games*. Cambridge, Mass.: MIT Press.
- HEINIMANN, F. (1987/1945): *Nomos Und Physis*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- HOBBS, T. (1651/1976): *Leviathan*. Frankfurt: Ullstein.
- JACOBSEN, H. J. (1996): "On the Foundations of Nash Equilibrium," *Economics and Philosophy*, 12, 67-88.
- KLIEMT, H. (1991): "Das Denken in Ordnungen Und Die Möglichkeiten Ordnungspolitischen Handelns," in *Ordnung Und Freiheit. Symposium Aus Anlaß Des 100. Jahrestages Des Geburtstages Von Walter Eucken Am 17. Januar 1991*, ed. by M. E. e. a. Streit. Tübingen: Mohr, 31-59.
- LAHNO, B. (1995): *Versprechen - Überlegungen Zu Einer Künstlichen Tugend*. München, Wien: R. Oldenbourg Verlag.
- LEWIS, D. (1975): *Konventionen*. Berlin: De Gruyter.
- MACKIE, J. L. (1980): *Hume's Moral Theory*. London: Routledge.
- MCKENZIE, R. B., and G. TULLOCK (1984): *Homo Oeconomicus*. Frankfurt: Campus.
- MECKLING, W. (1976): "Values and the Choice of the Model of the Individual in the Social Sciences," *Schweizerische Zeitschrift für Volkswirtschaft und Statistik*, 112/4, 545-565.
- MILGROM, P., and J. ROBERTS (1993): *Economics, Organization and Management*. Englewood Cliffs: Prentice Hall.
- MISES, L. V. (1940/80): *Nationalökonomie. Theorie Des Handelns Und Wirtschaftens*. München: Philosophia Verlag.
- MORGENSTERN, O. (1972): "Descriptive, Predictive and Normative Theory," *Kyklos*, 25, 699-714.
- MORGENSTERN, O., and G. SCHWÖDIAUER (1976): "Competition and Collusion in Bilateral Markets," *Zeitschrift für Nationalökonomie*, 36, 217-245.
- MYERSON, R. B. (1991): *Game Theory - Analysis of Conflict*. Cambridge, London: Harvard University Press.
- NÖLDEKE, G., and L. SAMUELSON (1997): "A Dynamic Model of Equilibrium Selection in Signaling Markets," *Journal of Economic Theory*, 73, 118-156.
- OFEK, H. (2001): *Second Nature*. Cambridge: Cambridge University Press.
- RAPHAEL, D.-D. (1969): "British Moralists," Oxford: Oxford University Press.
- SCHÄFER, H.-B., and C. OTT (2005): *Lehrbuch Der Ökonomischen Analyse Des Zivilrechts*. Berlin et al.: Springer.
- SHELLING, T. C. (1978): *Micromotives and Macrobehavior*. New York and London: Norton.
- SEN, A. K. (1973/1982): "Behaviour and the Concept of Preference," in *Choice, Welfare and Measurement*. Oxford: Basil Blackwell, 54-73.
- SIEGWART, G. (1997): "Explikation," in *Dialog Und System*, ed. by W. Löffler, and E. Runggaldier. Sankt Augustin: Academia, 15-45.
- SILLER, S. (1998): "Letter to the Editor," *Journal of Theoretical Biology*, 195, 413-417.
- SPENCE, M. (1973): "Job Market Signaling," *Quarterly Journal of Economics*, 87, 355-374.
- STRAWSON, P. F. (1962): "Freedom and Resentment," *Proceedings of the British Academy*, 187-211.

- STÜTZEL, W. (1981): "Grundtexte Zur Sozialen Marktwirtschaft. Zeugnisse Aus Zweihundert Jahren Ordnungspolitischer Diskussion," Stuttgart-New York: Fischer.
- WEISE, P., W. BRANDES, T. EGER, and M. KRAFT (2004): *Neue Mikroökonomie*. Heidelberg und New York: Physica-Verlag.
- ZAHAVI, A. (1975): "Mate Selection - a Selection for Handicap," *Journal of Theoretical Biology*, 53, 205-214.
- ZAHAVI, A., and A. ZAHAVI (1998): *Signale Der Verständigung - Das Handicap Princip*. München: Insel.