

Christian Meyer

## **Hyperkooperativität und Wir-Intentionalität**

Ein interaktionssoziologischer Blick auf  
Tomasellos Forschungen

### **Einleitung**

Michael Tomasellos Arbeiten haben in den letzten Jahren dazu beigetragen, an einen in der Soziologie seit ihren Anfängen diskutierten und nach wie vor rätselhaften Tatbestand zu erinnern und ihn aus Sicht der evolutionären Psychologie mit neuer Evidenz zu versehen. Der rätselhafte Tatbestand besteht darin, dass Menschen in der Lage sind, zu einem handelnden ‚Wir‘ zu verschmelzen, das über die Individuen hinausgeht, aus denen es besteht. Beispiele hierfür im sozialen Leben sind zahlreich: man denke an Rituale wie das gemeinschaftliche Gebet oder den gemeinsamen Gesang, an das kollektive Marschieren beim Militär, das gemeinsame Musizieren, etwa in einer Jam Session, oder an den Tangotanz zu zweit: in all diesen Aktivitäten verschmelzen Individuen zu einem Wir, dessen Handlungen nicht auf Handlungen der beteiligten Individuen zu reduzieren ist. Stattdessen verteilen sich erstens Handlungen und Intentionen auf die beteiligten Individuen, und erst durch ihr Zusammentreten resultieren erfolgreiche ‚Wir-Handlungen‘. Zweitens werden ‚Wir-Handlungen‘ während des Tuns durch die fortlaufende, wechselseitig responsive Moment-zu-Moment-Koordination, Abstimmung und Adjustierung individueller Beiträge erzeugt: sie entstehen erst durch ein emergentes Zusammenspiel der Aktionen der beteiligten Individuen. Dieser zweite Punkt ist wesentlich zum Verständnis von Wir-Handlungen, wird bei Tomasello jedoch weitgehend übersehen.

Tomasellos große Leistung besteht vielmehr darin, die primordiale Sozialnatur des Menschen in den Fokus einer allgemeinen anthropologischen Theorie sowohl in ihren evolutionären, als auch ihren kulturtheoretischen Bezügen gestellt zu haben. Er hat die Möglichkeit der Etablierung eines im ‚Wir-Modus‘ laufenden sozialen Handelns dem Menschen als Alleinstellungsmerkmal zugeschrieben (Tomasello 2008, S. 6) und zur Grundlage ei-

ner Theorie menschlicher Kultur und Gesellschaft gemacht. Seine Einsicht in das Vorhandensein einer primordialen Sozialnatur des Menschen bietet besonderes Potenzial für die Erklärung von gemeinschaftlichen, koordinierten Aktivitäten und verkörperten Interaktionen wie sie in jüngster Zeit in einem breiten Themenspektrum soziologisch erforscht werden (z.B. chirurgische Eingriffe, *computer supported cooperative work*, architektonische und Ingenieursplanung, u.a.).

Dieser Text diskutiert die Möglichkeiten, aber auch die Grenzen einer Orientierung interaktionsoziologischer Forschung an Tomasellos Modell am Beispiel von Forschungen zur verkörperten Interaktion im Sport. Nach einer Vorstellung der interaktionstheoretischen Positionen in Tomasellos Arbeiten werden zwei seiner inhaltlichen Konzepte bzw. Annahmen anhand von Daten über Wir-Handlungen im Sport empirisch zur Diskussion gestellt: (1) Die Erzeugung, Umsetzung und Charakteristik einer Wir-Intentionalität und (2) die Erzeugung und Handlungsrelevanz eines *common grounds* für Wir-Handlungen. Im Fazit wird auf dieser Basis eine Kritik an einigen Prämissen der Tomaselloschen Interaktionsforschung geübt sowie ein Vorschlag zur konstruktiven Weiterentwicklung seiner Position gemacht.

## Interaktionstheoretische Positionen in Tomasellos Forschungen

### Die Experimente

Zu den soziologisch bedeutsamen Erkenntnissen der evolutionspsychologischen Forschungen Michael Tomasellos gehört die Einsicht in die generell altruistische, kooperative und soziale menschliche Natur. Tomasello zeigt in seinen Experimentalstudien, dass Menschenkinder schon früh eindrucksvolle Zeugnisse ihrer Fähigkeit und Bereitwilligkeit zum Zusammenschluss zu einem handelnden ‚Wir‘ geben. In mehreren Experimenten halfen z.B. erst 18 Monate alte Kinder von sich aus Erwachsenen beim Bewältigen von deren Vorhaben, indem sie Gegenstände aufhoben, die die Erwachsenen fallen gelassen hatten und an die sie nicht ohne Weiteres wieder herankamen. Oder sie sekundierten unaufgefordert, indem sie den Erwachsenen Handlungsoptionen aufzeigten, die diese aus ihrer Perspektive nicht sehen konnten (Warneken/Tomasello 2006). Die Kinder waren also in der Lage, sowohl das Vorhaben (die Intention) der Erwachsenen zu identifizieren als auch deren Perspektive zu übernehmen und die Objektwelt aus ihrer Sicht zu sehen.

Menschenaffen kooperieren demgegenüber zwar auch, insbesondere mit Menschen (z.B. sind sie auch ohne Aussicht auf eine Belohnung hilfsbereit gegenüber ihren Pflegern), denken zumindest bei Futter aber in erster (und

offenbar einziger) Linie an sich selbst (Jensen et al. 2005). In einem Experiment mussten Schimpansen gleichzeitig an zwei Enden eines Seiles ziehen, um an ein mit Futter bestücktes Brett zu gelangen (Melis et al. 2006). Sie holten nur dann einen Artgenossen zu Hilfe in den Versuchsraum, wenn die Seilenden so weit voneinander entfernt waren, dass sie die Aufgabe nicht alleine lösen konnten. Dies zeugt zwar von großen kognitiven Fähigkeiten, aber eben nicht von sozialen bzw. kooperativen oder gar altruistischen Motiven. In den Analysen wird nicht näher darauf eingegangen bzw. bleibt nach wie vor rätselhaft, ob und, wenn ja, wie die Affen jeweils über ihr gemeinsames Ziel kommunizieren und sich in ihren Aktivitäten abstimmen, obwohl dies interaktionssoziologisch die zentrale Frage wäre.

Dies gilt auch für die Studien mit Menschenkindern. Eine Untersuchung von Warneken und Tomasello (2006), die Menschenkinder mit Schimpansen vergleicht, zeigt, dass Menschen ab drei Jahren eine moralische Verpflichtung gegenüber Kooperationspartnern verspürten, mit denen zusammen sie zuvor in einer Aufgabe ein ‚Wir‘ gebildet hatten. Hatte ein Kind dem anderen geholfen, eine Belohnung zu bekommen, so war die Neigung zur Gegenhilfe größer, als wenn die Belohnung zufällig gegeben wurde. Schimpansen zeigten demgegenüber in einem ähnlichen Experiment keinerlei Neigung, dem Gegenüber in einer kooperativen Situation eher zu helfen als in einer nichtkooperativen (vgl. Warneken et al. 2007).

An dieser Stelle muss allerdings zur Kenntnis genommen werden, dass Tomasellos Experimentdesigns (eine Anzahl an Videos der Experimente ist einsehbar auf: [http://www.eva.mpg.de/psycho/study-videos\\_de.php](http://www.eva.mpg.de/psycho/study-videos_de.php)) bis auf wenige Ausnahmen die soziale Situation auf eine Weise gestalten, dass das Handlungsprojekt und die damit verbundene Intention zunächst auf eine (menschliche, erwachsene) Person beschränkt ist. Das jeweilige Experiment beobachtet dann, ob das Projekt von einem zweiten Akteur (Menschenkind, Schimpanse) identifiziert und mit einer assistierenden Handlung unterstützt oder vervollständigt wird. Aus methodischen Gründen (Variablenisolierung, Ausschluss von Störvariablen) kommuniziert der erwachsene Mensch dabei nicht, sondern verfolgt eigensinnig sein Projekt, so dass auch so gut wie keine Abstimmung bzw. Interaktion über das Ziel selbst stattfindet. Vielmehr überlässt es der Experimentator komplett dem Interaktionspartner (Kind/Affe), sein Handlungsziel zu identifizieren, seine Handlung zu antizipieren und auf dieser Basis zu agieren. Die dabei zu Tage tretenden Fähigkeiten sind zwar in der Tat beeindruckend, aber als soziale Situationen erscheinen die Experimente dadurch merkwürdig und ungewöhnlich non-interaktiv und unpraktisch. Die methodisch begründete Reduktion der Interaktion auf rein mental ablaufende Antizipationsprozesse soll freilich die angeborene und primordiale Fähigkeit des Perspektivwechsels und der Handlungsantizipation elizitieren. Allerdings kann diese Fähigkeit auf Seiten der Kinder auch als Ergebnis

sozialer Spielerfahrungen und eines daher stammenden Interesses an sozialen Emergenzen verstanden werden. In der Tat wirken die Kinder in den Experimentvideos immer, als wollten sie mit ihren Aktivitäten den Experimentleiter zum sozialen Spiel auffordern.

Durch die methodische Reduktion auf die Identifikation der Intention von *alter* werden die üblichen, von der Soziologie vielfach beschriebenen sequenziellen interaktionalen Verfahren zur prozeduralen Herstellung von Intersubjektivität und zur Abstimmung (Schegloff 1992) ausgeblendet, obwohl Tomasello sie theoretisch an einigen Stellen (s.u.) für bedeutsam erklärt. Durch das Experimentdesign werden Intentionen von vorneherein als *Gegenstand* von Interaktion ausgeschlossen und als vorab vorliegende mentale Zustände (bzw. deren Qualitäten) festgelegt. Gerade im Spiel aber sind Intentionen selbst oft emergente Anteile und Objekte der Interaktion, wodurch das Spiel gerade ‚vom Hundertsten ins Tausendste‘ und ‚vom Hölzchen aufs Stöckchen kommt‘. Diese interaktiv-emergenten Dynamiken von Intentionalitäten in der Interaktion werden methodisch *a priori* ausgeschlossen.

## Die menschliche Hyperkooperativität

Tomasello wertet die menschliche Anlage zu der in den Experimenten demonstrierten, wie er sie nennt (2008, S. 172), „Hyperkooperativität“ als entscheidenden Faktor für die spezifisch menschliche kulturelle Evolution. Nach Tomasello ist die außergewöhnliche kognitive Infrastruktur des Menschen spezifisch aus kollaborativen Tätigkeiten (z.B. Jagen in der Gruppe, arbeitsteilige Werkzeugherstellung) und nicht etwa aus gestiegenen Anforderungen des Konkurrenzkampfes oder des Soziallebens erwachsen. Dies bezeugten sogar einige anatomische Eigenschaften des Menschen (z.B. der Frequenzbereich des Gehörs, der spezifisch auf die Stimmfrequenz der Artgenossen und nicht – wie etwa bei Eulen – auf diejenige von Beutetieren ausgerichtet ist). Ähnlich wie Goffman (1983) geht also auch Tomasello davon aus, dass die Partikularitäten der menschlichen Sozialität aus gemeinsam in einer kooperativen Haltung und mit gemeinsamem Ziel ausgeführtem praktischem Handeln resultieren. Sie werden dann *top down* von diesem gemeinsamen Projekt gesteuert (Tomasello 2008, S. 197). Dies steht z.B. im Kontrast zur Konversationsanalyse, die davon ausgeht, dass es gerade die von praktischen Zwängen entlasteten Situationen sind, in denen sich die Ursprungsform menschlicher Sozialität zeigt (vgl. Heritage 2008).

Alle für eine kulturell emergente, d.h. von ihren biologischen Wurzeln entkoppelte soziale Entwicklung notwendigen Konventionen, von einfachen geteilten Normen über Medien der symbolischen Kommunikation bis zu verstetigten Institutionen, basieren nach Tomasello auf der menschlichen

Veranlagung zur Hyperkooperativität – der Fähigkeit, die Intention von *alter* zu verstehen und seine Perspektive (kognitiv, normativ und emotional, d.h. empathisch) einzunehmen, der Befähigung, durch gemeinsame, geteilte Aufmerksamkeit vom Ich zum Wir zu gelangen und unter der Prämisse zu handeln, dass *alter* weiß, dass *ego* weiß, dass *alter* weiß, was das Wir vorhat, sowie, nicht zuletzt, einem Gefühl der moralischen Verpflichtung gegenüber einem solchen einmal etablierten ‚Wir‘.

Hyperkooperativität wird mit grundlegend prosozialen Motiven realisiert und umfasst somit gemeinsame Aufmerksamkeit und ein geteiltes Verständnis der vorliegenden Situation (Tomasello 2008, S. 108). Tomasello (2008, S. 107) weist aber auch darauf hin – auch wenn er dies empirisch nicht beachtet –, dass bei kooperativem Handeln wichtig ist, dass die Interaktionspartner – während sie die gemeinsame Absicht erfolgreicher Kommunikation produzieren – sich permanent miteinander abstimmen und aneinander anpassen müssen.

## Zur Bedeutung der Sinne

Für die menschliche Hyperkooperativität ist das Verfolgen der Blickrichtung von *alter*, mit dem auch dessen Aufmerksamkeitsfokus nachvollzogen werden kann, von fundamentaler Bedeutung (Tomasello/Carpenter 2007). Dem visuellen Sinn sind Tomasello (2008, S. 328) zufolge eine Reihe von Vorteilen eigen. Nichtmenschliche Primaten folgen den Blickrichtungen anderer über die Kopfhaltung und erlangen dadurch ein Verständnis von potentiell relevanten und interessanten Phänomenen in der Umwelt. Menschliche Kleinkinder hingegen folgen bereits sehr viel genauer den Augen selbst und damit der Blickrichtung ihrer Interaktionspartner, überdies schauen sie immer wieder auch zurück auf die Person, deren Blick sie folgen, so dass sie über die Identifikation von *alters* mutmaßlicher Wahrnehmung, seinen potentiell beobachteten Objekten und seinen körperlichen Bewegungen auf seine Intentionen schließen können (Tomasello 2008, S. 229).

In diesem Zusammenhang verweist Tomasello (2008, S. 196) auch auf die „interessante morphologische Tatsache“, dass Menschen als einzige Primaten eine großflächige weiße Sklera besitzen, die ihre jeweilige Blickrichtung von außen, d.h. für Dritte beobachtbar macht. Menschen, wie er es ausdrückt, tun dadurch ihre Blickrichtung öffentlich kund (*advertise*), was zur Folge hat, dass erstens „situations in which the individual may count on others using that information collaboratively or helpfully, not competitively or exploitatively“ in der menschlichen Evolution, Gesellschaft und Kulturgeschichte vorherrschen (Tomasello 2008, S. 197). Zweitens, und auf Interaktion bezogen, bedeutet dies für Tomasello, dass kooperativen Tätigkeiten

immer ein, wie er es nennt (2008, S. 94 f.), „rekursives“ Moment innewohnt: „when we start to act intending that our actions should be coordinated with – then we have to design our actions so that they are self-evidently perspicuous“ (Levinson 1995, S. 241, zit. in Tomasello 2008, S. 205). Auch wenn die Ähnlichkeit zum ethnomethodologischen Reflexivitätskonzept (*accountability*), auffällig ist<sup>1</sup>, nutzt Tomasello hier den Frege'schen Begriff der Rekursivität, der die Fähigkeit bezeichnet, selbst-referentielle Verfahren auszuführen. Durch diese Form der Selbstreflexivität und -referentialität, so Tomasello (2008, S. 106), entstehen aus Akten des Helfens und Teilens wechselseitige Erwartungen und Kooperationsnormen und aus dem Verstehen von Zielen und Intentionen gemeinsame Ziele und kommunikative Intentionen (Tomasello 2008, S. 215).

Für all dies, betont Tomasello (2008, 195, n. 1), sei der visuelle Sinn unabdingbar, da auditive Reize allen kopräsenten Personen simultan gesendet werden. Für die visuelle Sphäre hingegen gilt: „I see something, and to know whether you also see it I have to look at you (unlike audition). I also have to look at you to see if you see me seeing it, and so forth and so on. Thus, nocturnal animals with no vision (...) would never evolve joint attention“ (Tomasello 2008, S. 195, n. 1). Ebenfalls mit dem visuellen Sinn verbunden sind Zeigegesten, die von Kindern schon im Alter von neun Monaten kommunikativ eingesetzt werden, während Affen sie nur gegenüber Pflegern (zum Fordern von Futter), nicht jedoch untereinander nutzen (Tomasello 2006; Bullinger et al. 2011). Ferner basieren auch Techniken des Lernens, die über Imitation laufen, auf dem visuellen Sinn und der Fähigkeit, sich in das Gegenüber hinein zu versetzen. Während Affen keineswegs „nachäffen“ (Tomasello 1996), funktioniert die Kulturvermittlung unter Menschen häufig über Nachahmung.

Diese starke Betonung der kooperativen Vorteile der Visualität, die Tomasello mit der Schütz'schen wie der Goffman'schen Soziologie teilt, steht im direkten Gegensatz zum Sequenzialitätsprinzip, das die Ethnomethodologie, Konversationsanalyse und Systemtheorie betonen (Meyer, 2014). Letzteres hebt die intersubjektivitätssichernde Möglichkeit wechselseitig getakter verbaler Interaktion hervor (Schegloff 1992), die auf zwei Hauptbeobachtungen basiert: Erstens, dass in der Interaktion *ego* schon während der Äußerungen von *alter* Bestätigungssignale über deren Verstandenhaben geben kann, die von *alter* bereits simultan zur Äußerungsproduktion

---

1 Es proklamiert die Identität von Praktiken, die Bedeutung konstituieren, und Praktiken, die Bedeutung interpretieren (Garfinkel 1967, S. 77–79): „Phenomena of order are identical with procedures for their endogenous production and accountability“ (Garfinkel 2002, S. 72).

zur Kenntnis genommen werden können. Zweitens wird Intersubjektivität prozedural auch durch die Sequenzierung von Äußerungen selbst hervorgebracht: indem *ego* auf *alters* Äußerung mit einer Reaktion, die einen Interpretationsvorschlag beinhaltet, anschließen kann, die in einem dritten Turn wiederum von *alter* validiert oder korrigiert werden kann. Tomasellos pauschale Annahme ist demgegenüber, dass auch diese Form der verbalen sequenziellen Kooperation auf vorsprachlichen, visuell basierten Praktiken der Intentionalitätsattribution und des Perspektivwechsels aufruht.

### **Common ground und Projekt**

Die menschliche, auf dieser spezifischen Form von Intentionalitätsverstehen und Perspektivwechsel beruhende Veranlagung zur Kooperativität ist Tomasello zufolge primordial, d. h. stabil und auch etwa von Belohnungsreizen unabhängig. Aufgrund ihrer permanent wirksamen Fähigkeit, die Intentionen von *alter* zu verstehen, und der Möglichkeit, dieses Verständnis etwa durch Assistenz und Mitwirkung auch mitzuteilen, haben Menschen die Gabe, ein geteiltes Aktivitätsprojekt zu etablieren.<sup>2</sup>

Für solche Projekte gilt generell, dass je mehr unter den Interaktanten geteilt ist, desto weniger explizit gemacht werden muss, um erfolgreich zu kooperieren (vgl. Tomasello 2008, S. 79, S. 100). Denn je mehr gemeinsames Wissen wir zu teilen annehmen können, umso einfacher ist es, unseren Interaktionspartnern Intentionen zuzuschreiben und ihre Handlungen zu antizipieren. Ein wechselseitig ratifiziertes ‚Projekt‘ trägt also als *common ground* dazu bei, dass Kommunikationsleistungen (z. B. das Zeigen auf einen Gegenstand) auch in ihrer Komplexität reduziert werden (der Gegenstand wird nicht in Bezug auf seine Form, Farbe oder Funktion zur Kenntnis genommen, sondern z. B. als potentielles Versteck verstanden). Tomasellos Vorstellungen weisen durchaus Ähnlichkeiten zu etablierten soziologischen Modellen wie diejenigen von – um nur zwei zu nennen – Schütz oder Goffman auf. Schütz geht davon aus, dass die Lebenswelt der Menschen ohnehin sinnhaft vorkonstruiert ist, so dass bereits kulturell – z. B. über Typisierungen – eine ständige Reduktion von Deutungsoptionen stattfindet (vgl. Schütz 1974). Die Vertrautheit der Lebenswelt stellt auch die Ressourcen bereit, damit in konkreten Situationen das Zuschreiben von Motiven verschiedener

---

2 Ein Beispiel aus Tomasellos Experimenten ist ein Spiel, bei dem Gegenstände versteckt werden. Die ebenfalls zur Intersubjektivität beitragende interaktionale Sequenzialität, die Aktivitäten dieser Art aufweisen, wird von Tomasello weiter nicht beachtet.

Art adäquat geschehen kann. Da der Sinn von Handlungen nach Schütz im subjektiven, der Handlung selbst vorausgehenden Handlungsentwurf liegt, der dem oder der Handelnden jeweils eine Vorstellung vom Ziel und Endpunkt der Handlung verleiht, kann er durch Motivzuschreibungen von *alter* identifiziert werden.

Goffman hat mit seinen Konzepten des *working consensus* (1959, S. 9 f.) und später des *frames* (1974) ebenfalls darauf aufmerksam gemacht, dass soziale Situationen über Bedeutungszuschreibungen vordefiniert sind, die manchmal allein über den Kontext der Aktivität konstituiert werden, oft implizit bleiben, stets allerdings fragil und immer wieder verhandelbar sind. Mit dem Eingehen (oder zumindest Nicht-Anfechten) eines Arbeitskonsenses bzw. einer Rahmung verständigen sich die Teilnehmer über ihre jeweilige Aufgaben und Konzentrationsrichtungen sowie über zur gegenwärtigen Aufgabe gehörende oder von dieser abweichende Aktivitäten: die ‚Hauptaktivität‘, mögliche ‚Nebenaktivitäten‘ und eventuelles zu ignorierendes ‚unwillkürliches Verhalten‘. Alle können unterschiedlich definiert sein und müssen immer wieder in Übereinstimmung gebracht und an den geltenden Rahmen angepasst werden, wodurch dieser ebenfalls aktualisiert wird.

## Kollektive Intentionalität

Tomasello (2008, S. 6, S. 181; 2009, S. XIII; 2014, S. 3; Tomasello et al. 2005, S. 6 f.) bezieht sich nicht auf die soziologische Tradition, sondern nachhaltig auf Theorien kollektiven Handelns aus der Philosophie des Geistes, namentlich auf Margret Gilbert, Michael Bratman, John Searle und Raimo Tuomela (vgl. dazu Schmid/Schweikard 2009). Sie verstehen das Vorliegen einer Intentionalität als zentrales Kriterium des Handelns, und Tomasello folgt ihnen darin. Intentionalität bestimmen sie als eine Qualität mentaler (innerer) Zustände: als Ausgerichtet-Sein auf ein gemeinsames, in der Zukunft liegendes Handlungsziel, das gegeben sein muss, bevor eine Handlung ausgeführt wird (vgl. Searle 1983, S. 1). Liegt der entsprechende Zustand bei mehreren Personen zugleich vor und wissen diese diesbezüglich voneinander, dann sprechen die genannten Theoretiker von geteilten oder gemeinsamen Intentionen (bzw. – Tuomela folgend – „Wir“-Intentionen) inklusive einem (nun Gilbert folgend) gemeinsamen Ausgerichtet-Sein (*joint commitment*) auf ein gemeinsames Ziel (*joint goal*). Gemeinsame Intentionen sind durch Prozesse der gemeinsamen Aufmerksamkeit sowie durch wechselseitig existierendes Wissen strukturiert, was wiederum zum wechselseitigen Beobachten der Aufmerksamkeit des je anderen führt (Tomasello 2008, S. 181; Tomasello 2009, S. XIII–XIV). Von Bratman übernimmt Tomasello (vgl. Tomasello et al. 2005, S. 6) in leichter Modifikation drei Wesensmerkmale von gemein-

samen kooperativen Aktivitäten: (1) dass die Interaktanten wechselseitig responsiv zueinander sind, (2) dass ein gemeinsames Ziel in dem Sinne existiert, dass jeder Teilnehmer das Ziel hat, dass das Wir in gegenseitiger Kenntnis zusammen X tut, und (3) dass die Teilnehmer ihre Handlungspläne und Intentionen koordinieren, was erfordert, dass sie alle Seiten der Interaktion verstehen (also Rollentausch vornehmen können) und sich dadurch potentiell untereinander helfen können.

Auf Searle wiederum bezieht sich Tomasello (2009, S. 58 f.) besonders in seinen Thesen zur Generalisierung von Einzelsituationen gemeinsamen Handelns hin zu Formen kollektiver Intentionalität und schließlich zum Entstehen von sozialen Institutionen, Normen und Kultur, die sämtlich nur existierten, weil die Menschen an sie glaubten und auf der Basis dieses Glaubens handelten.<sup>3</sup>

## Kooperation und Kultur

Die intersubjektive Verfasstheit von sozialen Aktivitäten mag auch der Grund sein, dass Menschen mit Freude kooperieren (Tomasello/Carpenter 2007, S. 123), während Schimpansen ihren Partner eher instrumentell einsetzen, *alter* und *ego* mithin nicht zu einem Wir verschmelzen. Die Emergenzsprünge von der individuellen Aufmerksamkeit auf *alters* Aufmerksamkeit hin zur geteilten Aufmerksamkeit vor dem Hintergrund einer gemeinsamen Situationsdefinition, von der individuellen Manipulation *alters, ego* bei der Erfüllung einer Aufgabe zu helfen, hin zur kooperativen Kooperation im Rahmen eines gemeinsamen, normativ besetzten Wir-Projekts sowie von der Aktivität einer Gruppe, bei der sich Einzelaktivitäten summarisch aggregieren, hin zu einer kollaborativen Aktivität, bei der die Summe mehr ist als seine Bestandteile, geschehen mithin nur beim Menschen (Tomasello/Carpenter 2007). Nicht das kantische Verständnis von Raum, Zeit, Menge und Kausalität oder ihre operative Intelligenz unterscheiden die menschliche Spezies von ihren nächsten Verwandten, den Schimpansen, die ihr immerhin genetisch ähnlicher sind, als es Schimpansen zu anderen Primaten sind. Denn in Bezug auf diese Fähigkeiten ähneln sich Schimpansen und zweijährige Menschen. Es ist Tomasello zufolge vielmehr die soziale, kooperative Kompetenz, spezifisch das Erkennen von Intentionen und das Hineinversetzen in *alter* zu kumulieren, wechselseitig zu verschränken und so auf der Basis

---

<sup>3</sup> Tomasello kritisiert Searles Intentionalitätstheorie allein dahingehend, dass dieser geteilte Intentionen auch Hyänen zugesteht, die gemeinsam handeln, und nicht nur Menschen (Tomasello/Rakoscy 2003, S. 133 n. 5).

geteilter Situationsdefinitionen miteinander in einem Wir zu kommunizieren und zu handeln, die das Spezifikum menschlicher Interaktionalität und Sozialität ausmacht – ein Tatbestand, der immerhin zu den faszinierendsten Gründungsmotiven der Soziologie seit Cooley, Mead und Schütz zählt.

Tomasello bezieht sich auch auf diese soziologischen Denktraditionen so gut wie gar nicht, greift aber auf die Sozialvertrags-Kontroversen des 17. und 18. Jahrhunderts zurück. Dabei gibt er Rousseau und seiner Theorie des von Natur aus hilfsbereiten und kooperativen Menschen, der erst von der Gesellschaft korrumpiert wird, Recht gegenüber Hobbes und dessen Annahme des von Geburt aus egoistischen und unkooperativen Menschen, der erst von der Gesellschaft eines Besseren belehrt wird (Tomasello 2009, S. 3). Allerdings betont Tomasello, dass im Allgemeinen jede menschliche Kultur auf der basal hilfsbereiten, mitteilbaren und großzügigen Ausstattung, mit der Menschen geboren werden, aufbaut. Wenn Menschen älter und unabhängiger werden, lernen sie, selektiver mit ihrem altruistischen Verhalten umzugehen (Tomasello 2009, S. 44 f.).

Levinson (1995, S. 227), der sich stattdessen auf Hume und Machiavelli bezieht, betont, dass die von Tomasello herausgearbeiteten Fähigkeiten zur sozialen Kooperation (die er mit Hume assoziiert) sehr viel anspruchsvoller seien als eine „agonistische Intelligenz“ nach Machiavelli: „*Machiavellian intelligence* is child's play, a lower-order computational ability; *Humeian intelligence* (coordination through implicit contract) is the adult stuff“ (Herv. im Orig.).

Am Anfang der menschlichen Kulturentwicklung steht also die Wir-Intentionalität. Da kooperierende Gruppen erfolgreicher waren als Einzelgänger, hatten sie einen evolutionären Vorteil, der sich zugleich ständig verstärkte. Hinzu kommt, dass sie durch ihre kooperativen Fähigkeiten mittels ihrer Sprache, aber auch anderer Instrumente (Werkzeuge, Artefakte, etc.) die Möglichkeit erlangten, ihr Wissen zu externalisieren, zu objektivieren und auf Dauer zu stellen, was Affen nicht gelingt. Die Fähigkeit zur kulturellen Weitergabe ist also nach Tomasello den Menschen biologisch gegeben – so lautet das Fazit seiner Version einer naturalistischen, nicht-reduktionistischen Theorie der menschlichen Gattung.

### **Hyperkooperativität und Wir-Intentionalität in der sozialen Interaktion: empirische Schlaglichter aus dem Sport**

Tomasello hat zu Recht darauf aufmerksam gemacht, dass die kooperative Grundanlage des Menschen vor allem im Rahmen gemeinsamer Aktivitäten relevant und sichtbar wird. Seine Hypothese ist zugleich auch phylogenetisch: er geht davon aus, dass diese Infrastruktur sich in solchen Aktivitäten (z. B. dem Jagen) herausgebildet hat. Aus diesem Grund erscheint es sinnvoll,

einen ähnlichen Typus von Aktivitäten auch zur Grundlage einer näheren empirischen Betrachtung der Stichhaltigkeit seiner Hypothesen zu wählen. Hierfür eignet sich professioneller Sport, bei dem sich mehrere Individuen mit einem Ziel zusammenschließen und gemeinsame Handlungen ausführen, um dieses Ziel zu erreichen.<sup>4</sup> Hyperkooperativität misst sich an der Fähigkeit zum erfolgreichen Ausüben von Wir-Handlungen, d.h. von gemeinsamen Handlungen, die im Wir-Modus ausgeführt werden. Im Folgenden soll daher an zwei Beispielen gezeigt werden, dass das Teilen von Intentionen als Ressource nicht ausreicht, um den tatsächlichen Verlauf von kooperativen Aktivitäten zu erklären: (1) der Erzeugung, Umsetzung und Charakteristik einer Wir-Intentionalität und (2) der Erzeugung und Handlungsrelevanz eines *common grounds* für Wir-Handlungen.

### **(1) Erzeugung, Umsetzung und Charakteristik von Handlungen im Wir-Modus**

Als Beispiel für Handlungen im Wir-Modus werden an dieser Stelle Spiele der Teamsportart Handball vorgestellt. Im Handball schließen sich mehrere Personen mit dem Ziel zusammen, ein Spiel gegen ein gegnerisches Team zu gewinnen. Dabei übernimmt jeder Spieler eine Rolle (Position in Abwehr und Angriff), die – unter Einbezug des Regelsystems – im Zusammenschluss ein sinnvolles System zur Verteidigung des eigenen Tors und zum Erzielen eines Tors beim Gegner ergibt. Es handelt sich um ein System verteilter Aufgaben, das, wenn sich jeder auf seine Aufgabe konzentriert, dazu führt, dass die Verteidigung schwer zu überwinden ist und der Angriff zum Torerfolg kommt. Das Zusammenspiel der Einzelaktionen wird von einer Bundesligamannschaft freilich im Training jeden Tag geübt, sowohl in der Abwehr als auch im Angriff.

---

4 Die folgenden Beispiele stammen aus dem Forschungsprojekt „Kommunikation unter Druck – Praktiken der Verständigung im Spitzensport“ an der Universität Bielefeld (Projektleitung: Jörg Bergmann & Christian Meyer an der Fakultät für Soziologie; Klaus Cachay & Carmen Borggreve an der Fakultät für Psychologie und Sportwissenschaften). Die empirischen Beispiele wurden von Ulrich v. Wedelstaedt erhoben und zusammen mit ihm analysiert. Die Transkripte und Standbilder wurden von Eva Fenn erstellt. Beiden möchte ich an dieser Stelle herzlich danken. Alle Beispiele stammen jeweils aus dem professionellen Leistungssport. Im Fall der Handballmannschaft handelt es sich um Teams aus der deutschen Bundesliga. Die in der Fallstudie Boxen begleiteten Sportler sind zu einem Großteil Sportsoldaten, -polizisten oder in vergleichbaren Fördermaßnahmen. Obwohl es sich bei der Sportart also um Amateurboxen handelt, sind die aller meisten der hier Beteiligten professionelle Vollzeitathleten bzw. -trainer. Die im Folgenden dargestellten Aufnahmen entstanden bei einem nationalen Meisterschaftswettkampf. Alle Transkripte sind anonymisiert. Eine Übersicht über die verwendeten Transkriptionszeichen findet sich im Anhang.

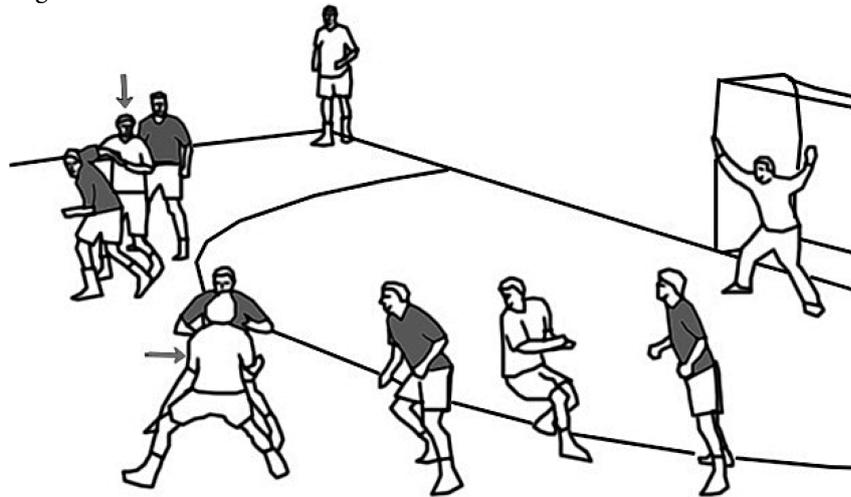
Für den Angriff – aus dem das folgende erste Transkript stammt – werden Spielzüge eingeübt, die das gegnerische Abwehrsystem verwirren und vor Koordinationsprobleme stellen sollen, so dass ein Torerfolg wahrscheinlicher wird. In der Regel wird der geplante Spielzug zu Beginn des Angriffs vom ‚Spielmacher‘ (meist in der Rückraum-Mitte-Position) angesagt oder durch ein gestisches Zeichen angekündigt. Er legt somit neben der generellen „Wir“-Intention, gewinnen zu wollen, auch den intendierten spezifischen Weg zum Torerfolg vorab fest, so dass tatsächlich davon ausgegangen werden kann, dass das Handeln der Gruppe von einer Intention angeleitet wird, die – da der Spielzug im Training immer wieder geübt wurde – einen relativ klaren Inhalt besitzt. Inwieweit der Inhalt mental-repräsentationaler oder nicht vielmehr verkörpert-kinetischer Art ist, ist eine andere Frage.

Vor ein größeres Problem sieht sich die gemeinsame Intention allerdings direkt in der Folge gestellt, wenn die eigene Mannschaft auf die Abwehr der Gegner trifft. Nun stellen sich Kontingenzen vieler Art ein: Akte der Gegner und die starke Abwehr im Allgemeinen stören den eigenen Plan, nicht-antizipierte, unkonventionelle Abwehrleistungen machen das noch so gut ausgeklügelte Projekt zunichte, es muss aufgrund von eigenen Fehlern oder Unzulänglichkeiten der Mitspieler revidiert werden, das Material streikt (etwa ein rutschiger Boden, ein aufgrund zu vielen Harzes allzu klebriger Ball), der Zuschauerlärm oder Musik in der Halle lenken ab. Das bedeutet, dass Spielzüge stets in die ‚Hier-und-Jetztigkeit‘ der aktuellen, von Widrigkeiten der genannten Art geprägten Situation implementiert werden müssen. Sie können so gut wie nie eins zu eins umgesetzt werden, sondern müssen angepasst werden.

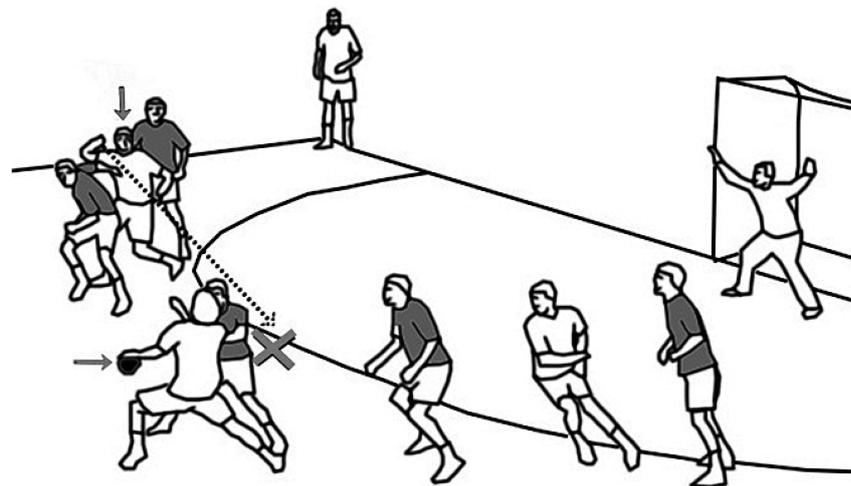
Für jede soziale Situation gilt daher aus Sicht der Teilnehmer die Eigenschaft der ‚Einzigartigkeit‘, die Garfinkel (1996, S. 7 u. ö.) mit dem der scholastischen Philosophie entstammenden Begriff der *Haeccitas* gefasst hat. Garfinkel geht davon aus, dass Teilnehmer soziale Situationen gewissermaßen immer wieder aufs Neue zum ersten Mal bewältigen müssen („each next first time“, vgl. Rawls 2005, S. 184). Dies liegt darin begründet, dass die Teilnehmer stets vor dem ‚Applikationsproblem‘ (Wolff 1976) stehen. Der abstrakte, im Training eingeübte Plan des Spielzugs muss kreativ und responsiv an die vorliegende Situation angepasst werden. Dabei entstehen große Spielräume, die Chancen und Risiken bieten. Die einzigen Ressourcen intersituationaler Kontinuität und Stabilität von Praktiken sind damit die verkörpert-reflexive, responsiv-*accountable* wechselseitige Verständigung der Teilnehmer zur Interpretation der Situation und zum gemeinsamen Handeln. Auch die Anpassung an die konkrete Situation selbst erfordert – da es sich um komplexe Handlungsabläufe mehrerer Individuen handelt – die minutiöse Abstimmung von Projektmodifikationen und die spontane Revision vorgesehener Pläne sowie das gemeinsam abgestimmte beherzte Ergreifen von Opportunitäten.

Das erste Beispiel zeigt eine solche Situation, in der spontan eine Gelegenheit von zwei Angriffsspielern genutzt wird. Dabei macht der Kreisläufer (senkrechter Pfeil) dem ballbesitzenden rechten Rückraumspieler (waage-rechter Pfeil) ein Zeichen (Bild 2), wo er hinlaufen wird. Der Rückraumspieler antizipiert die Bewegung (Bild 4) und spielt den Ball (schwarzer Kreis) im richtigen Moment über den Abwehrspieler (Bild 5), der Kreisläufer bekommt ihn und es kommt zum Torerfolg (Bild 6).

Transkript 1: Demonstration eines ‚Strebens‘ und antizipatives Zuspiel im Angriff

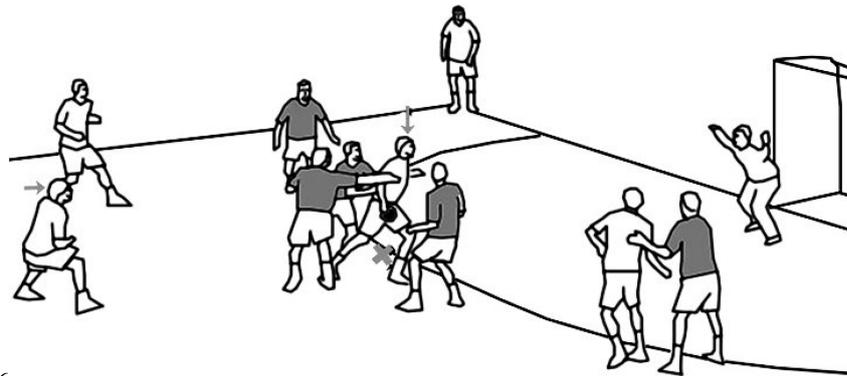


1



2





6

Das Geschehen passiert extrem schnell: es dauert nur zwei Sekunden. In der Videoaufzeichnung wird zudem deutlich, wie koordiniert und flüssig die Bewegungen der beiden Spieler sind: während der Kreisläufer sich von seinem Abwehrspieler losreißt und mit Blick und Gesten seine Laufrichtung anzeigt (die Gesten seiner Hände sind zugleich Bewegungen des Strebens), beginnt der Rückraumspieler bereits mit der Ausholbewegung des linken Armes, mit dem er den Blindpass dann dahin spielt, wohin der Kreisläufer gelaufen sein wird, wenn der Ball ankommt (Bilder 2 bis 4; die Stelle, wohin der Ball gespielt wird, ist mit einem X gekennzeichnet). Es ist also keineswegs so, dass einer der beiden Spieler die Aktion vorbereitet und steuert und dabei abwartet, bis der andere soweit ist, seinen Teil des Plans auszuführen. Vielmehr bildet die Bewegung beider Spieler von Anfang an zusammengenommen eine orchestrierte Gesamthandlung, mit der die spontane Gelegenheit genutzt wird, dass sich eine Abwehrlücke aufgetan hat, die dem Kreisläufer eine Gelegenheit zum Torwurf bietet.

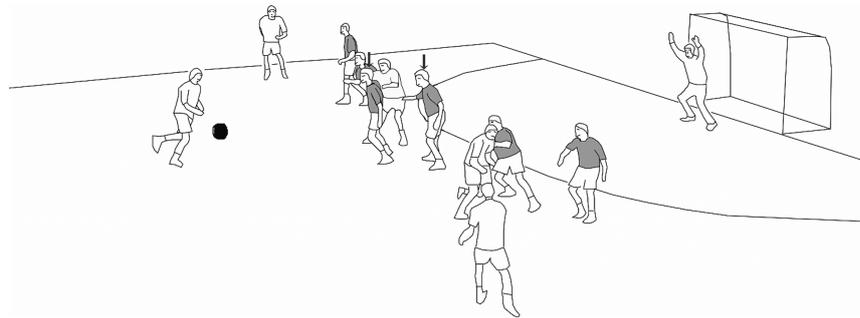
Während also in der Tat eine Wir-Intention zum Spielgewinn im Allgemeinen und zum Torerfolg im Besonderen vorhanden gewesen sein mag, so hat sich hier eine spontane Gelegenheit zum Torwurf ergeben, die in dieser Form und Detaillierung nicht vorab intentional repräsentiert war. Die Intention (Torerfolg) wurde vielmehr permanent anhand der aktuellen Situation responsiv zu Mitspielern, Gegnern und Umwelt revidiert und rekonfiguriert. Die Intentionalität der konkreten Wir-Handlung ist an dieser Stelle emergent und nicht vorab festgelegt. Sie lag in zu keinem Zeitpunkt in einer mentalrepräsentationalen Form vor, sondern allenfalls in einer Form der „verkörperten Intentionalität“. Loenhoff (2012, S. 189) bestimmt einen verkörperten Begriff von Intentionalität unter Rückgriff auf Gehlen als „Erwartungsvorgriff einer jeden gerichteten Bewegung auf Erfolg, Antwort und Rückschlag“. Bereits einfache gesteuerte Bewegungen antizipierten „ihren Umgangserfolg“ jenseits einer mentalen Intentionalität. Im Anschluss an Merleau-Ponty schreibt er sämtlichen Funktionen des Leibes, da sie mit dem Vollzug von

Bewegungen einhergehen, den Status einer präreflexiven Intentionalität zu. Die Bewegungserfahrung stelle damit als leibliche Intentionalität einen ursprünglichen und eigenständigen Modus des Zugangs zur Welt dar (Loenhoff 2012, 192). Das mit dieser Intentionalität verbundene verkörperte Wissen über (durchaus interaktionale) „Bewegungsgestalten“ wird, wie auch im vorliegenden Fall, permanent mit Widerstandserfahrungen der materiellen Umwelt konfrontiert (Loenhoff 2012, S. 196; vgl. auch Loenhoff 2010, S. 62–64).

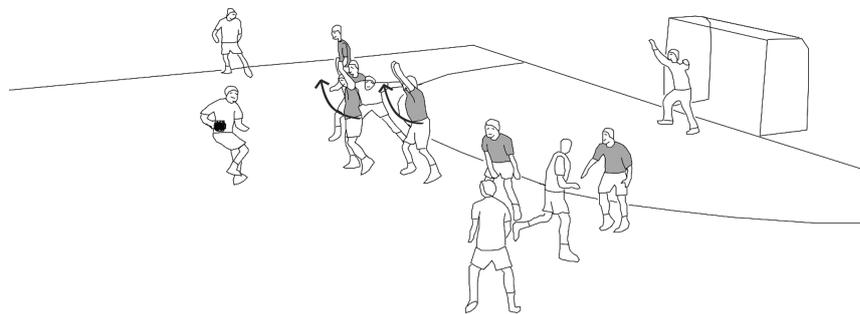
Das nächste Beispiel betrifft ebenfalls eine Form der ‚verkörperten Intentionalität‘, die in der sportlichen Interaktion zur Anwendung kommt. In der Abwehrarbeit ist es von grundlegender Bedeutung, dass jeweils zwei Personen (im folgenden Transkript durch je zwei Pfeile gekennzeichnet) einen Angreifer stoppen, der eine Person allein leicht wegdrücken, umspielen oder umwerfen könnte, oder dass sie gemeinsam einen Wurf blocken. Die Bewegungen, einen Zweier-Abwehrblock zu bilden, müssen dabei wohlkoordiniert sein, obwohl die Aufmerksamkeit beider Spieler auf den ballbesitzenden Angreifer ausgerichtet ist, und nicht auf den Abwehrblock-Partner. Dies gelingt in der Praxis erstaunlich gut: das Verhalten von Abwehrblocks ist durch eine eindrucksvolle Synchronität gekennzeichnet.

#### Transkript 2: Synchrones Handeln in der Abwehr

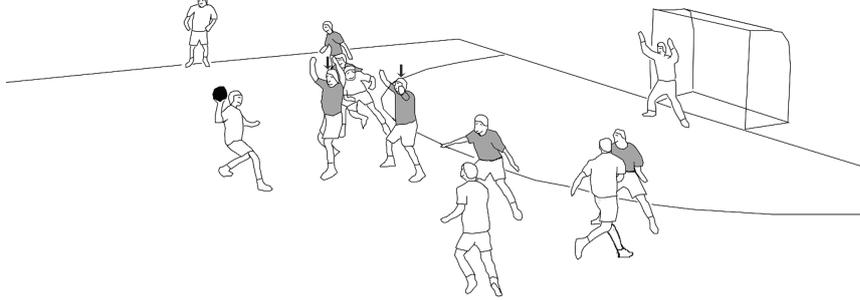
1



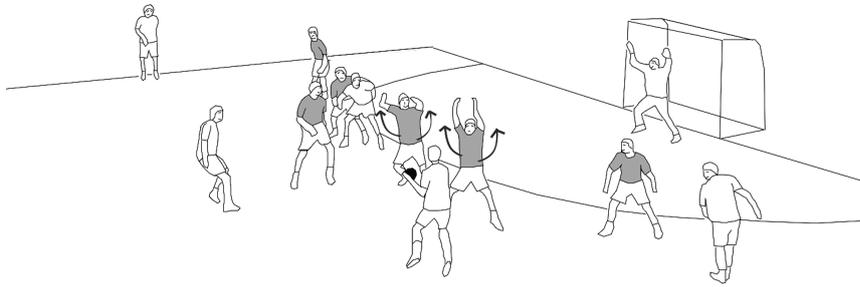
2



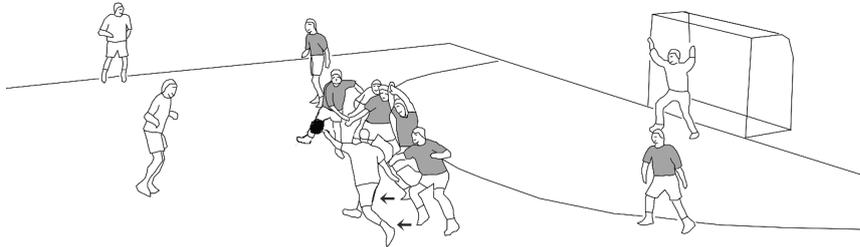
3



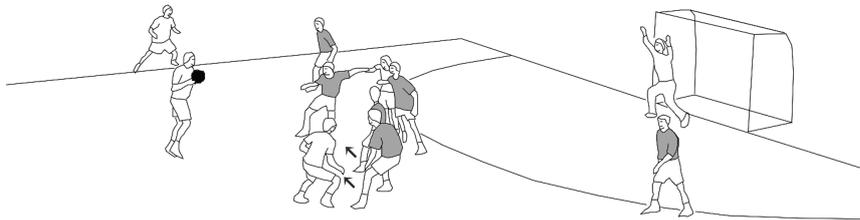
4



5



6



In den Bilderreihen 1 bis 3 und 4 bis 6 ist jeweils sehr gut zu erkennen, wie ein Zweier-Abwehrblock exakt synchron in seinen Bewegungen den Angreifer davon abhält, einen Sprungwurf oder ein Zuspiel an den Kreisläufer zu starten. Erst als es den Angreifern gelingt, diese Synchronität zu durchbrechen, gelingt auch ein Wurf auf das Tor. Auch hier ist deutlich, dass die Abwehrspieler – wenngleich eine allgemeine unspezifische Wir-Intention (etwa: „kein Tor Zulassen“) vorliegen mag – die Details des Wir-Handelns jeweils spezifisch an die sich ständig ändernde Situation anpassen, dabei aber unbedingt im Wir-Modus, d.h. in synchronen Bewegungsabläufen, bleiben müssen. Die Ressourcen, die sie dazu nutzen, bestehen zum einen aus im Training tausendfach eingeübten Bewegungsabläufen, die über die periphere Wahrnehmung koordiniert werden, während die zentrale Wahrnehmung auf den Gegner ausgerichtet ist; zum anderen, da der gegnerische Angriff immer wieder neue Herausforderungen bietet, bestehen sie aus über Schlagwort-Rufe ablaufenden Absprachen *in situ* („jetzt“, „los“, „raus“, „nimm/geh du“, „hab’ ich“, etc.).

Sowohl im Angriff als auch in der Abwehr ist das handelnde Wir ein hochgradig fragiles Gebilde, das ständig durch abnehmende Konzentration und Aufmerksamkeit zu zerfallen droht. Es muss aus diesem Grund immerzu aktualisiert werden, was über verschiedene Instrumente – robusten Körperkontakt untereinander, Motivationsrufe, Ansprachen und Rituale in Time-Outs – erreicht wird.

Beide Beispiele zeigen, dass ein handelndes Wir zwar vermutlich über eine, wenn auch nicht permanent mental präsente, allgemeine Intention verfügt („Gewinnen“, „alles Geben“, „zum Torerfolg Kommen“, „kein Tor Zulassen“), die konkreten, minutiös koordinierten tatsächlich ablaufenden Handlungen aber keinesfalls erschöpfend aus dieser erklärt werden können, sondern weitere Ressourcen erfordern – insbesondere Responsivität, Antizipation, Routinen, auf *alter* erweiterte Bewegungsschemata. Sie gilt es im Rahmen einer Forschung über die menschliche Kooperationsfähigkeit zu erforschen, denn sie stellen die komplexere verkörpert-kognitive Aufgabe dar, dergegenüber eine *a priori* vorliegende Wir-Intentionalität eine Banalität ist. Dies wird auch in den nun folgenden Beispielen aus dem Boxen deutlich.

## **(2) Erzeugung und Handlungsrelevanz eines common grounds für Wir-Handlungen<sup>5</sup>**

Obwohl das Boxen auf den ersten Blick eine Individualsportart ist, werden auch hier in sehr ausgeprägter Form Wir-Handlungen ausgeführt. Die

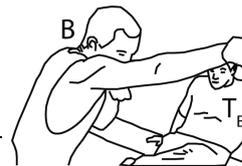
---

<sup>5</sup> Im Folgenden beziehe ich mich ausführlich auf unsere Analyse in Meyer/v. Wedelstaedt (2015); vgl. auch Meyer/v. Wedelstaedt 2013.

Existenz eines *common ground* ist im Boxen ebenso wichtig wie im Handball. Er wird in der Vorbereitungsphase und damit in den Wochen, Monaten oder sogar Jahren vor einem Kampf etabliert, in denen Boxer und Trainer zusammenarbeiten. Am Wettkampftag findet zwischen Boxer und Trainer eine erneute und das Gegebene aktualisierende umfassende Einstimmung aufeinander statt. Deren wichtigstes Element ist die Phase der Vorbereitung in der Kabine, die bis unmittelbar vor Beginn des Kampfes andauert. Für den Trainer ist diese unmittelbare Vorbereitungsphase auch deshalb von besonderer Bedeutung, weil er sich in der Regel mit mehreren Boxern an einem Wettbewerb beteiligt und sich durch sie jeweils auf den einzelnen einstellen kann. Das folgende Transkript zeigt vier Ausschnitte aus der Vorbereitung in der Kabine, die insgesamt circa 45 Minuten dauert. Sie ist gekennzeichnet von Aktivitäts- und Ruhephasen. Dem späteren Rhythmus des Boxkampfes entsprechend wechseln Boxer und Trainer dabei meistens zwischen drei Minuten Vorbereitung und einer Minute Regenerationsphase.

**Transkript 3: Vorbereitung auf den Boxkampf in der Kabine<sup>6</sup>**

01 T und GLEICH wieder drÜck nach VORne führn  
 02 (1,0)  
 03 T ja::  
 04 (2,4) \_\_\_\_\_  
 05 T schön agressiv ihn DRÜCKn



((...))

06 T nich nich ins RENnen kommen  
 07 ( ) auffpassn auf det ding \_\_\_\_\_  
 08 =und dann gehste wElter (-) ja  
 09 schön über die FÜHrungshand drückn  
 10 =KUCKn nochmal  
 11 und rechnen das was kommt (.)  
 12 weglassen  
 13 =sofort wieder übernehmen  
 14 =GEGEN schlagn  
 15 oder=oder ohne HANDlung übernehmn  
 16 =den druck IMmer noch versuchn nach VORne zu führn  
 17 IHN anne WAND stelln KEvin (--)  
 18 ja=der muss GLEICH von anfang merken wat LOS is



<sup>6</sup> T = Trainer, blau (in den Abbildungen T<sub>B</sub>); B = blauer Boxer.

((...))



19 T nach LINKS (.) HIER (.) RAUS (.)

20 ho::=da

21 das de HIER gleich RÜber trittst

22 (4,0)

23 T ja::=da stehste auch GUT (.)

24 (notfalls) HA:ken (.)

25 oder die GRAde zum KOPP

26 (3,5)

27 T FÜHrungshand

28 =FÜHrungshand=RÜber

29 treten=sobald de sein

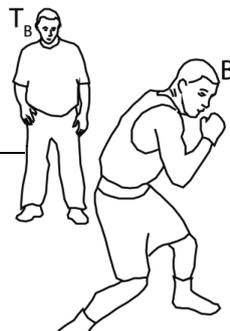
30 ding SIEHST

31 (2,0)



32 T geNAU::=geNAU: so: KEvin

((...))



33 T un überNEHmen (-)

34 jaWOLL

35 =un weiter nach VORne

36 bam BAM (--)

37 so is SCHÖ:N (--)

38 HOB (1,4) hob (--)

39 hob (-) hob (--)

40 FÜHrungshand

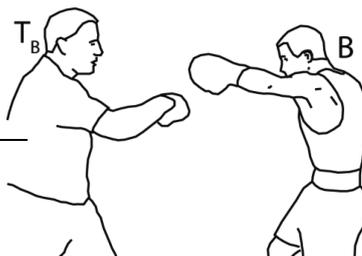
41 (6.0)

42 T jaWO:LL (---)

43 so is RICHTig

44 =TSCHAK (---)

45 jaWO:LL



Ein entscheidendes Element in der Abstimmung der beiden Körper aufeinander bildet die körperliche Symmetrisierung. Sie wird in der Vorbereitungsphase des Kampfes in der Stellung der beiden Körper zueinander sichtbar. Zu Beginn sitzt der Trainer auf einer Bank in der Kabine, beobachtet den Boxer in seiner Vorbereitung und spricht ihn gelegentlich an (Zeile 04, Standbild). Anschließend steht er auf und positioniert sich seitlich zum Boxer, so dass beide einander von vorne sehen können (Zeile 07, Standbild). Während der Trainer zu Beginn in relativ langen, elaborierten Sätzen mit dem Boxer über die erwartete Kampfsituationen (vor allem Zeilen 06-11) spricht, sind seine verbalen Äußerungen jetzt bereits stark verkürzt und richten sich immer stärker an den simultan vollführten Bewegungen aus bzw. treten sogar hinter diese zurück. An diesem Punkt nimmt die Vorbereitung eine Form des *Call-and-Response* ein: Der Trainer simuliert einen Schlagrhythmus, einen Bewegungsablauf oder ein Angriffsmuster und der Boxer vollzieht dies – möglichst gleichförmig – nach.

In der nächsten Stufe dieses Prozesses der Symmetrisierung parallelisieren Trainer und Boxer ihre visuellen Perspektiven: Der Trainer vollzieht ein Angriffsmuster, wobei ihm der Boxer von hinten über die Schulter schaut (Zeile 19, Standbild). Er gibt dabei auch detaillierte Anweisungen zur Positionierung im Raum (Zeilen 19, 21 & 23) und verdeutlicht diese gestisch und durch seine Blickrichtung (in diesem Fall durch Geste und Blick auf den Boden: Zeile 21, Standbild). Anschließend führt der Boxer das gleiche Bewegungsmuster aus, wozu er vom Trainer an exakt die Stelle des Raumes dirigiert wird, an der dieser selbst zuvor stand (Zeile 23, Standbild). Anschließend begibt sich der Trainer in die Position, die der Boxer vorher eingenommen hat, und schaut diesem seinerseits über die Schulter (Zeile 32, Standbild).

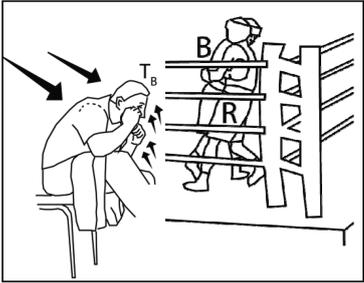
Die verbalen Äußerungen des Trainers beschränken sich in dieser Phase, besonders, wenn er die Bewegungsmuster des Boxers kommentiert, auf wenige Worte, die in einer besonderen Betonung und Sprachmelodie vorgebracht werden (z. B. das stark betonte „genau“, Zeile 32). Exakt derartige Äußerungen – wie im Folgenden noch gezeigt wird – verwendet der Trainer auch im Verlauf des Kampfes zur Anleitung des Boxers. Aufgrund der extremen Schnelligkeit des Boxgeschehens und der extrem limitierten Aufnahmefähigkeit des Boxers während des Kampfes (vor allem im Angesicht schwerer physischer Gewalt, die seinem Körper widerfährt), eignen sich lange elaborierte Einwüfe für den Kampf nicht.

In dieser Weise findet zwischen Trainer und Boxer also ein Synchronisationsprozess in Hinsicht auf die Zeitlichkeitsstruktur statt, mit der sie auch in der nachfolgenden Kampfsituation konfrontiert sein werden. So werden Trainer und Athlet in die Lage versetzt, auch unter diesen extremen Bedingungen auf koordinierte Weise Anweisungen zu geben und zu empfangen.

Sprachlich-kinästhetische Synchronisation und körperlich-taktile Symmetrisierung zwischen Boxer und Trainer finden in der Vorbereitungsphase ihren Höhepunkt im so genannten ‚Einschlagen‘ (Zeilen 33–45). Dabei schlägt der Boxer in die ‚Pratzen‘ (offene Handschuhe, die eine Schlagfläche bieten) des Trainers. Der Trainer simuliert hierbei zum Teil den späteren Gegner, indem er etwa bestimmte Ausweichmanöver vollführt oder auch ‚zurückschlägt‘ (dann aber verhältnismäßig leicht), bis die Einnahme der Boxstellung<sup>7</sup> erreicht ist. Zum anderen gibt der Trainer auch Instruktionen, die der Boxer umsetzen soll, allerdings – weil diese Phase die Temporalstruktur des späteren Kampfes, also die extreme Schnelllebigkeit, bereits vorwegnimmt – nur noch in Form kürzester Interjektionen (z. B. „hob“, Zeilen 38–39). Die so aufeinander abgestimmten Trainer und Boxer begeben sich anschließend in den Kampf, der im Folgenden betrachtet wird.

Während des Kampfes sitzt der Trainer (zusammen mit einem Assistenten) wenige Meter vom Geschehen entfernt am Rande des Rings. Er beobachtet den Kampf permanent und ruft fast ohne Unterbrechung Anweisungen in Richtung seines Sportlers. In der Situation des Kampfes wird die Ausübung einer Wir-Handlung besonders gut sichtbar. Das folgende Transkript zeigt zwei kurze Ausschnitte aus der ersten und zweiten Runde eines Kampfes.

#### Transkript 4: Boxer und Trainer im Kampf<sup>8</sup>

<p>01 T jaWO::LL          02 (1)          03 ru:hi:ch KEvin          03 NICH mit ma:cht          04 (1,3)          05 ru:hich          06 WIEder GRAde KEvin (-)          07 ANbietn (-)          08 JAwo:ll (0,5)          09 schö::n KEvin (1)</p> <p>((...))</p>		
---	---	--

7 Das ist die Grundstellung der Boxer, in der sie Kopf und Oberkörper mit Händen bzw. Oberarmen vor gegnerischen Schlägen schützen.

8 T = Trainer, blau (in den Abbildungen T<sub>B</sub>); B = blauer Boxer; A = Assistent, blau (neben dem Trainer sitzend, in den Abbildungen nicht sichtbar); R = roter Boxer.

10 T ruHI (.)

11 ja bleib DRANne

12 =bleib DRANne KEvin

13 =je jaWO:LL (-)

14 un WIEder

15 A ja

16 T ruHICH am MANne (-)

17 ZU machn am MANN

18 A ei:NE:

19 (1,4)

20 T KEvin

21 (0,9)

22 T KEvin

23 am mann ZU machn

24 DREISsig KEvin

25 (1,6)

26 T je:: ((klatscht))

27 ey:: ((klatscht))

Im Transkript wird erkennbar, wie die vorherige Synchronisation und Symmetrisierung der Körper in den Kampf hinein wirkt: Im ersten Abschnitt des Transkripts nimmt der Trainer mehrfach selbst die Boxstellung ein – also das Schützen von Kopf und Oberkörper durch Hände und Arme (Zeile 03, Standbild) – und führt das ‚Tauchen‘ aus. Das heißt, er vollzieht ein Ausweichmanöver, bei dem er den Oberkörper zunächst schnell vorbeugt und dann (seitlich versetzt) wieder in die Ausgangsposition zurück bringt. Den Boxern dient das Tauchen dazu, sich unter die Schläge des Gegners zu bewegen und außerhalb der Schlagreichweite wieder nach oben zu kommen. Im zweiten Transkriptabschnitt vollzieht der Trainer den gesamten Bewegungsablauf des Tauchens noch einmal besonders schnell (Zeilen 13, linkes Standbild, & 16, Standbild). Die durch den Trainer hier vollzogenen Bewegungen korrespondieren auf das Genaueste mit dem Kampfgeschehen, das er in diesem Moment im Ring beobachtet (der Boxer befindet sich ebenfalls in einer Tauchbewegung: 13, rechtes Standbild). Der Trainer vollzieht die Bewegungen, die der Boxer in diesem Moment ebenfalls vollzieht oder vollziehen sollte.

Im Transkript wird außerdem erkennbar, wie der Trainer seine verbalen Einwüfe meist auf einzelne Wörter beschränkt. Die Interjektionen des Trainers – z. B. „schön“ (Zeile 09), das mehrfach wiederholte „jawohl“ (Zeilen 01,

08 & 13) oder auch die (zum Teil in besondere Lautstärke erfolgende, Zeile 20) Nennung des Namens des Athleten (Zeilen 03, 06, 09, 12, 20, 22 & 24) – erfahren handlungspraktische Implikation über ihre spezifische prosodische Ausgestaltung (Lautstärke, Betonung, Rhythmik o.ä.). Die enorme verbale Verkürzung erfährt über diese zusätzliche Ausgestaltung eine Kontingenzreduktion und gewinnt an Eindringlichkeit. Nur über diese Form der Beitragsgestaltung wird eine Kompatibilität mit dem Kampfgeschehen und der hier vorherrschenden besonderen Schnellebigkeit der Ereignisse erreicht.

Der Trainer verfolgt den Kampf aber – wie wir gesehen haben – nicht nur über die visuelle Wahrnehmung der Situation oder die verbale Begleitung der Aktionen des Boxers, er vollzieht den Kampf auch körperlich selbst mit. Interessant ist in diesem Zusammenhang die mehrfache Verwendung des Auspruchs „ruhig“ (Zeilen 03, 05, 10 & 16) und dessen körperliche Begleitung. An zwei Stellen ruft der Trainer „ruhig“, als Athlet und mimetisch er selbst am Ende eines Ausweichmanövers sind (Zeilen 03 & 16, Standbilder). An der verbalen Begleitung der Rückkehr in die Ausgangsposition wird deutlich, dass der Trainer hier simultan einer Eskalation seiner eigenen körperlichen Performanz entgegensteuert (so dass er an dieser Stelle nicht weiter mimetisch ‚mitkämpft‘, um seiner genuinen Aufgabe – der verbalen Anleitung des Boxers – angemessen nachkommen zu können) und sich zugleich an den Boxer wendet. Er spricht mit dem „ruhig“ auch ihn an, ist bemüht, ihn aus der unmittelbaren Nahkampfsituation herauszuwinden, die Übersicht über das Kampfgeschehen wiederherzustellen und eine neue Angriffsbewegung aus der Distanz entstehen zu lassen. Das „ruhig“ adressiert – anders gesagt – das Wir.

Ein weiteres mimetisches ‚Mitkämpfen‘ des Trainers neben dem Boxring hätte in diesem Moment seine Fähigkeit eingeschränkt, den Kampf zu verfolgen und Anweisungen zu geben. Gleichzeitig wird jedoch auch deutlich, dass der Trainer nur durch das mimetische ‚Mitkämpfen‘, also den körperlichen Nach- oder Mitvollzug des Geschehens im Ring, in der Lage ist, seine verbalen Anweisungen an die syntaktische und semantische Struktur des Kampfgeschehens anzupassen. Es ist generell auffällig, wie stark die körperliche Aktivität des Trainers mit der körperlichen Aktivität des Boxers im Ring korrespondiert. Besonders in Nahkampfphasen, in denen sich die Ausgangslage für taktische Manöver innerhalb von Sekundenbruchteilen ändern kann, bedient sich der Trainer einer Vielzahl von extrem kurzen, prosodisch und sprachmelodisch ausgestalteten Interjektionen. Eine andere Technik ist das Klatschen (Zeilen 26 & 27), das Bezug auf zuvor eingeübte Schlagrhythmen nimmt und dadurch handlungspraktische Bedeutung entfaltet. In diesem Spannungsfeld aus der Notwendigkeit zum körperlichen Mitvollzug und den hieraus entstehenden Nachteilen bewegt sich der Trainer während des Kampfes.

Der Trainer vollzieht also die Bewegungen nicht nur *ex post* nach, sondern er vollzieht das Kampfgeschehen simultan mit. Er antizipiert Aktionen des Gegners, reagiert in seiner eigenen Körperlichkeit auf diese und liefert auf dieser Grundlage auditive, handlungspraktische Anweisungen an den Boxer. Obwohl der Trainer den Schlägen selbst nicht ausgesetzt ist, sondern nur den Boxer beobachtet, der im Ring steht, wurde durch die Phase der Einrichtung der beiden Körper aufeinander während der Vorbereitung auf den Kampf – der zeitlichen Synchronisations- und körperlichen Symmetrisierungsarbeit – eine so enge Verbindung zwischen Boxer und Trainer als *common ground* etabliert, dass hier von sozialem Handeln im Wir-Modus gesprochen werden kann. Dies ergibt sich vor allem aus der unmittelbaren Verknüpfung der praktischen, verkörperten Wissensbestände der beiden Beteiligten und der Einrichtung der (para-)sprachlichen Muster von Äußerungen und Rezeption aufeinander (vgl. dazu auch Meyer/v. Wedelstaedt 2013, 86 ff.).

Anders als im Handball kommt es im Boxen nicht zur Bildung eines „Wir“ aus einer Vielzahl von Beteiligten, die eine Aufgabe gemeinsam bewältigen, die von einer Einzelperson nicht geleistet werden könnte. Gerade im Boxen wird im Allgemeinen der Kampfausgang der körperlichen und mentalen Leistungsfähigkeit des einzelnen Kämpfers zugeschrieben. Wie aber gezeigt wurde, findet auch hier eine unmittelbare und sehr enge Form der Wir-Kooperation statt, wenngleich diese, anders als im Handball, nicht dem erfolgreichen Ausführen einer gemeinsamen körperlichen Aktivität dient, sondern dem erfolgreichen Umsetzen von taktischen *in situ*-Anleitungen des Trainers durch den Athleten. Im Kampf selbst wird das Wir durch konstante Zurufe des Trainers an den Boxer aufrechterhalten, während der Trainer selbst die Situation des Boxers mimetisch mitvollzieht, um situativ angemessene Anweisungen in den Ring rufen zu können. Evident wird die Bedeutung dieser Verbindung in den Fällen, in denen eine Störung zwischen den Beteiligten auftritt (ein Beispiel bei Meyer/v. Wedelstaedt 2013, 90 f.).

Das vor allem mimetisch ausgerichtete Ritual des Einschlagens vor dem Kampf dient somit der Etablierung einer synchronisierten und symmetrisierten Wir-Körperlichkeit, die als *common ground* für den Kampf erforderlich ist. Dadurch findet zum einen eine Synchronisierung statt, die den mimetischen Mitvollzug des Kampfes durch den Trainer ermöglicht, zum anderen dient es als Basis für die Verständigung mittels extrem verkürzter Äußerungen (Klatschen, Schlagworte etc.) im Kampf. Wichtig in all diesen Beispielen ist ferner nicht so sehr das Blicken und Sehen des Blickens von *alter*, sondern die Nutzung des gesamten Sinnesspektrums sowie der Sequenzialität von Akten und Re-Akten.

## Der intentionalistische Hintergrund Tomasellos: ein Reparaturversuch

Die intentionalitätstheoretischen Ressourcen, auf die sich Tomasello mit seinem Ansatz beruft, sind generell nicht unumstritten, und viele der Kritikpunkte gehen in eine Richtung, die auch die hier vorgelegte empirische Evidenz nahelegt. Dreyfus (1993) etwa kritisiert Searle (1983) aus einer Heideggerianischer Perspektive: Intentionen seien keine stabilen, das gesamte Handeln über präsenten inneren mentalen Zustände, sondern würden erst dann explizit und relevant, wenn Störungen im Handeln auftauchten, etwa wenn im gemeinsamen Handeln implizit vereinbarte (bzw. routinehaft erwartbare) Assistenzleistungen von *alter* „unzuhanden“ würden. Im Normalfall müssten Intentionen für ein erfolgreiches Handeln bzw. eine gelingende Praxis gar nicht explizit vorhanden und schon gar nicht dauerhaft präsent sein.<sup>9</sup>

Daran schließt sich ein zweites Problem des intentionalistischen Ansatzes zur Erklärung gemeinsamen Handelns an, dass nämlich eine Auffassung von Intentionen als mentalen Zuständen (vgl. Schaubild Tomasello et al. 2005, S. 7) zu statisch ist, um das Phänomen gemeinsamer Intentionalität, wie es sich in Interaktionen präsentiert, zu fassen. Wie Gibbs (2001, 122) feststellt, werden viele Aspekte intentionaler Bedeutung und Handlung zumindest teilweise erst in dynamischen sozialen Interaktionen erzeugt und existieren keineswegs vorab als mentale Repräsentationen im Innern isolierter Individuen. Intentionen werden in Interaktionen überhaupt erst als solche herausgebildet, etwa wenn eine Person eine Äußerung tätigt, die dann von *alter* interpretiert und so in ihrer Bedeutung determiniert wird, was wiederum von *ego* ratifiziert oder bestritten wird. Die Zuschreibung von Intentionen durch *alter* ist oft wichtiger für die interaktionale Konstitution von gemeinsamen Intentionen als mögliche originäre Intentionen von *ego*. Z.B. wenn *ego* zufällig und geistesabwesend auf etwas blickt, was von *alter* beobachtet und zum

---

9 Tomasello gesteht diesen Punkt sogar zu – allerdings ohne ihn als Kritik an seinem Intentionalitätsmodell oder auch nur Experimentdesign zu verstehen –, wenn er darauf hinweist, dass Wir-Intentionalität (oder geteilte Intentionalität) besonders in „bumpy interactions in which things thought to be shared turn out not to be“ Tomasello (2008, S. 335 f.) explizit und reflexiv würden. Menschen werden sich ihrer geteilten lebensweltlichen Annahmen (um es in der Sprache der Sozialphänomenologie auszudrücken) und ihrer gemeinsamen Intentionen also dann bewusst, wenn sie fehlen, so dass gemeinsames Handeln ins Stocken oder zum Zusammenbruch kommt. Entgegen vieler sozialpragmatischer und sozialphänomenologischer Studien, die dazu vorliegen, fügt Tomasello hinzu, dass „the actual data for this hypothesis are not so numerous“ (Tomasello 2008, S. 335 f.).

Anlass genommen wird, dieses Etwas zum Thema und Gegenstand weiterer Interaktionen zu machen, und *ego* seiner Interpretation letztlich – aus welchen Gründen auch immer – nicht widersprechen möchte oder eine mögliche eigene Intention ihm erst in diesem Moment als solche bewusst wird. Solche emergenten Intentionalitätszuschreibungen sind keineswegs arbiträr und individuell-partikular, sondern basieren auf kulturell bereit gestellten Zuschreibungsvokabularien (vgl. Mills 1940). Soziologisch gesehen ist es also keineswegs so, dass Intentionen Handlungen immer als Ursachen vorausgehen und Handlungen als Wirkungen von Intentionen zu verstehen sind, sondern oft umgekehrt, dass Intentionen sozialen Praktiken als kommunikative Rationalisierungen, Legitimierungen oder Plausibilisierungen *ex post facto* hinzugefügt werden (vgl. Gibbs 2001, S. 117 f.).

Gibbs (2001, S. 113) gibt das Beispiel des Gesellschaftstanzes: wenn man als Paar tanzt, dann werden die eigenen Bewegungen von den Bewegungen des Partners mitgestaltet, während sie selbst zugleich die Bewegungen des Partners beeinflussen. Beide Bewegungsfolgen ko-evoluieren in einer nicht vorhersehbaren und nicht auf die Vorab-Intentionalitäten der Beteiligten reduzierbaren Weise. Dass nicht nur gemeinsames, geteiltes, sondern auch kollektives Handeln am besten als emergentes Produkt des Zusammenspiels von Gehirnen, Körpern und Maschinen zu verstehen ist, zeigt Hutchins' Studie (1995) über das Navigieren eines Flugzeugträgers der US-Marine. Jedes Mitglied der Schiffsbesatzung reagiert dabei jeweils adaptiv auf Umgebungsbedingungen, für deren Kontrolle es zuständig ist. Die einzelnen Handlungen jedes Besatzungsmitglieds können zwar möglicherweise unter Rückgriff auf seine Intentionen verstanden werden, aber es existiert keine kollektive Intention, aus der die richtige Bewegungsrichtung des Schiffs resultiert. Vielmehr tragen die Reaktionen jedes Einzelnen dazu bei, Aspekte im gemeinsamen Steuerungsprozess zu verändern, die wiederum Reaktionen bei den anderen Crew-Mitgliedern hervorrufen. Insgesamt stellt der gesamte Prozess ein umweltlich erweitertes Verfahren dar, bei dem mehrere Akteure, Routinen und Artefakte (z.B. nautische Datenschieber) zusammenspielen, um das komplexe Problem der Schiffsnavigation zu bewältigen. Das kollektive Handeln der Mannschaft entsteht erst aus sozialen Interaktionen zwischen Menschen in Verbindung mit der Umwelt, und nicht durch ausdrückliche Vereinbarungen zwischen den Besatzungsmitgliedern, wie das Schiff zu bewegen ist. Zwar mag eine geteilte Intention darüber vorhanden sein, dass das Schiff in den Hafen navigiert wird, nicht jedoch wie und mit welchen Arbeitsschritten dies vollzogen werden soll.

Die schiere Existenz gemeinsamer Intentionalität ist somit nicht ausreichend zur Erklärung von sozialen Handlungen und Praktiken, wie sie sich in ihrem aktuellen Verlauf darstellen. Intentionen sind immer unterdeterminierend in Bezug auf das situierte Handeln, das sich im Rahmen leiblicher

Beziehungen zur Umgebung vollzieht (vgl. Joas 1996). Eine Erklärung von sozialem Handeln kann sich daher nicht damit begnügen, nach Absichten, Motiven oder Plänen zu suchen, die – etwa im anschließenden Interview – ein Handeln kommunikativ mittels Motivzuschreibungen erklären (vgl. Renn 2006, S. 23).

Der zentrale Fehler der Erklärung von gemeinsamem Handeln aus gemeinsamen Intentionen, der gerade in den Tomaselloschen Experimenten sichtbar wird, aber eigentlich seinem Grundanliegen – der Demonstration der intrinsisch kooperativen und sozialen Natur des Menschen – auch widerspricht, liegt darin, dass sie von Subjekten als autonomen und isolierten Einheiten ausgeht, die einander beobachten und sich auf dieser Basis mental ineinander hineinversetzen (vgl. Breyer 2013), dabei jedoch nicht miteinander in interaktional responsiven Kontakt treten, aufeinander reagieren, einander berühren und hierdurch gemeinsam handeln – z. B. beim gemeinsamen Tragen eines Tisches (vgl. Shotter 2011). Gemeinsame Intentionen findet man nicht durch intellektuelle Übungen im empathischen Perspektivwechseln und mentale epistemische Operationen, sondern sie entstehen in sozialen Praktiken als dynamisch-emergente Korrelate von Interaktionen.

Der einzige Weg, Zugang zu den Intentionen von *alter* zu bekommen, besteht also darin, mit ihm im Rahmen praktischen Tuns zu interagieren, so dass sowohl Verstehen als auch Missverstehen prozedural sichtbar und sequenziell weiter verhandelt wird (vgl. Slaby 2014). Ein Konzept, das dies umsetzt und daher zur Beschreibung der praktisch-verkörpernten Dimension von Kooperation geeignet erscheint, ist „participatory sense-making“ (De Jaegher/Di Paolo 2007): es betont die gelebte Wechselseitigkeit der sozialen Existenz. Aus dieser als „enaktiv“ bezeichneten Perspektive ist Kognition (und damit auch Intentionalität) „embodied action“ (De Jaegher/Di Paolo 2007, S. 486). „Enaktive Intersubjektivität“ umfasst „wechselseitige Inkorporation“, bei der die Leiber beider Teilnehmer sich aufeinander ausweiten und eine gemeinsame „Interkorporalität“ bilden (Fuchs/De Jaegher 2009, S. 465). Dieses Konzept (frz.: ‚intercorporalité‘, im Deutschen auch ‚Zwischenleiblichkeit‘), auf das an dieser Stelle leider nicht ausführlich eingegangen werden kann, geht auf Merleau-Ponty zurück, z. B. wenn er erläutert, wie die Erfahrung der reflexiven Körpererfahrung als Basis für leibliche Intersubjektivierungsprozesse beim Kontakt mit *alter* dient:

„Meine rechte Hand wohnte dem Auftreten der aktiven Berührung meiner linken Hand bei. Es ist nichts anderes, wenn sich der Körper des Anderen vor mir belebt, wenn ich die Hand eines anderen Menschen drücke, oder wenn ich sie einfach betrachte. Indem ich erfahre, daß mein Leib ein ‚empfindendes Ding‘ ist, daß er reizbar ist – er und nicht nur mein ‚Bewußtsein‘ –, bin ich darauf vorbereitet zu verstehen, daß es andere Animalia und möglicherweise andere Menschen gibt.

[...] Wenn mir das Dasein eines Anderen dadurch evident ist, daß ich ihm die Hand drücke, so deshalb, weil sie sich an die Stelle der linken Hand setzt, weil mein Leib sich dem des Anderen durch jene ‚Art der Reflexion‘ einverleibt, deren Sitz er paradoxerweise ist. Meine beiden Hände sind ‚kompräsent‘ oder ‚koexistent‘, weil sie die Hände eines einzigen Leibes sind: Der Andere erscheint durch eine Ausdehnung dieser ‚Kompräsent‘, er und ich sind wie die Organe einer einzigen Zwischenleiblichkeit (*intercorporéité*).“ (Merleau-Ponty 2007, S. 245-46; Herv. im Orig.)

Es bildet sich bei der enaktiven Intersubjektivität also gewissermaßen ein gemeinsamer Handlungskörper, der sich nicht auf die beteiligten Individuen und ihre Intentionen reduzieren lässt. Das Handlungssubjekt entsteht durch die „moment-to-moment interaction of two subjects“ und wird durch körperliche Resonanz, emotionale Ein- und Abstimmung, die Koordination von Gesten und weitere Praktiken erreicht (Fuchs/De Jaegher 2009, S. 466).

Fuchs und De Jaegher (2009, S. 474) illustrieren die wechselseitige Inkorporation und das „participatory sense-making“ mit dem Beispiel der Vorwärtssteuerung in einem Tennisspiel, bei dem ein Return es erfordert, dass der Spieler den Flugweg des ankommenden Balls schon während des Schlags seines Gegenspielers ‚inkorporiert‘ und so in seinen Bewegungen vorwegnehmen kann:

„As a skilled tennis player, I not only incorporate the ball and its trajectory but also my opponent’s position, posture and movements. I feel the thrust and direction of his stroke as well as the momentum the ball receives, and with this, my own body’s reaction is already being prepared. Here, my lived body is also in an ambiguous state, fluctuating between the incorporated body of the other and my own embodied position. In a fluent phase of the game, even before one player strikes the ball, the other’s reaction unfolds, and this already influences the first player’s initial action. As this goes on reciprocally, both players are connected in a feedback/feed-forward cycle, and there are no gaps of reaction time. Thus, they both experience the holistic development of the situation which is co-constituted by their bodily movements. However, if one of the players makes a surprising or feinting move, the coordination breaks. Here, the mutual incorporation does not match up the partners exactly, but each meets an autonomous response. So the game consists of an oscillation between matches and mismatches, of in-phase and phase-delayed states.“

Diese Form der interkorporalen Koordination erfordert und erlaubt zu keinem Zeitpunkt mentale und intentionale Repräsentation.

Heidegger hat mit seinem Hammer-Beispiel bereits darauf hingewiesen, dass unser Leib sich beim Benutzen von ‚Zeug‘ auf das Instrument ausweitet und mit ihm verschmilzt (oder naturalistisch ausgedrückt: sein Körperschema in seiner kognitiven Repräsentation darauf ausweitet, vgl. Dotov et al. 2010),

so lange es nicht durch Beschädigung „unzuhanden“ wird. Ein ähnlicher Prozess findet in der Interkorporalität mit dem Leib von *alter* statt, mit dem Unterschied, dass sein Leib zugleich den Leib von *ego* inkorporiert, so dass die oben genannte doppelte Inkorporation stattfindet, die – wie im obigen Tennisbeispiel deutlich wird – jedoch fragiler ist, als die Verschmelzung mit einem Hammer (oder einem Tennisschläger), da sie auch doppelt kontingent ist.

Die Ablehnung mental-repräsentationalischer Grundlegungen gemeinsamen Handelns, die dieser Ansatz vollzieht, sowie seine Annahme, dass menschliche Leiber sich in der beschriebenen Form miteinander über körperlich-responsive Praktiken, rhythmische Synchronisierungen und emotionale Abstimmungsformen verschränken, stellt sich jeder Form der konzeptuellen oder ontologischen Reduktion von Wollen und Handeln auf isolierte individuelle Körper entgegen (Froese/Fuchs 2012, S. 214). Tomasello bietet das Potenzial einer solchen naturalistischen, aber nicht individual-reduktionistischen Begründung sozialen Handelns (und weist auch bisweilen darauf hin), schöpft sie aber aufgrund der fehlenden Basierung auf den genannten Theorien und der Orientierung seiner Experimentaldesigns an intentionalistischen Modellen nicht aus.

## Fazit

Entgegen der philosophischen Diskussion ist die entscheidende Frage zur Beziehung zwischen der menschlichen Veranlagung zur Hyperkooperativität und der situierten interaktionalen Praxis offenbar nicht, ob die der letzteren zugrunde liegenden Intentionen individuell und zu aggregieren oder kollektiv und emergent sind, sondern ob Intentionen überhaupt zur Erklärung von gemeinsamem Handeln notwendig sind. Denn wenn Prosozialität und Hyperkooperativität angeboren sind, dann ist erstens vielmehr die *Abwesenheit* von Kooperativität und nicht ihr *Auftreten* erklärungsbedürftig (Girke/Meyer 2011; Breyer 2013). Zweitens ist der Status von Intentionen als mentalen Zuständen, die zu realisierten Praktiken hinzukommen müssen, damit sie als gemeinsame Praktiken oder Handlungen bezeichnet werden können, unklar. Ist „Gewinnen Wollen“ oder „Helfen Wollen“ ein mentaler Zustand oder eine sozial und kulturell vermittelte Kategorie? Wer steht am Ursprung dieses mentalen Zustands? Tomasellos Einsichten in die altruistische und prosoziale Natur des Menschen würden nahelegen, dass gemeinsames Handeln als Selbstzweck zu verstehen ist, der – da er der menschlichen Natur entspricht – *per se* sozial motiviert ist. Das Vorhandensein spezifischer Intentionen wäre damit als zusätzliche Bedingung nicht notwendig. Die spezifische Intention ‚Gewinnen Wollen‘ im Handballspiel z.B. fiele damit in

eins mit der primordialen prosozialen Motivation des Menschen. Intentionen könnten dann allenfalls als Fokussierungselemente verstanden werden, die zwischen einer allgemeinen Prosozialität und konkretem gemeinsamem Handeln vermitteln, das sich daraus aber eben – wie gezeigt – nicht erschöpfend erklären lässt. Tomasello hat sich hier in ein logisches Problem verstrickt, das bereits Garfinkel bei Parsons, Durkheim bei Spencer und Hume bei Rousseau, Locke und Hobbes gesehen haben: wenn soziales Handeln aus sozial motivierten Intentionen resultiert, dann bleibt unerklärt, wie soziale Intentionen überhaupt zustande kommen. Sie müssen letztlich bereits auf einer vor-intentionalen Sozialität aufruhend, die der Wir-Intentionalität vorausgeht und die Bedingung ihrer Möglichkeit darstellt. Hume, der dieses Argument gegen die Sozialvertragstheoretiker angeführt hat, sollte daher der Gewährsmann einer Wir-Handlungstheorie sein. Sein Beispiel ist dann auch das des gemeinsamen Ruderns, bei dem keiner der Ruderer dem anderen ein explizites Versprechen gegeben hat, gemeinsam zu rudern. Gemeinsames Handeln ist eine *Praxis*:

„Two men, who pull the oars of a boat, do it by an agreement or convention, tho' they have never given promises to each other. Nor is the rule concerning the stability of possession the less deriv'd from human conventions, that it arises gradually, and acquires force by a slow progression, and by our repeated experience of the inconveniences of transgressing it. On the contrary, this experience assures us still more, that the sense of interest has become common to all our fellows, and gives us confidence of the future regularity of their conduct; and it is only on the expectation of this that our moderation and abstinence are founded. In like manner are languages establish'd by human conventions without any promise“ (Hume [1739] 1992, S. 490).

Auch wenn einige der philosophischen Gewährsleute Tomasellos konstatieren, dass „the social sciences presently are at a fairly primitive and undeveloped stage“ (Tuomela/Miller 1988, S. 369), bleiben Tomasellos interaktionstheoretische Erkenntnisse ihrerseits aus sozialtheoretischer Sicht unterreflektiert. Es liegt mittlerweile eine ganze Reihe von interaktions- und sozialtheoretischen Entwürfen aus pragmatistischer, phänomenologischer und praxistheoretischer Sicht vor, die angesichts der hier aufgezeigten Schwächen adäquatere Modelle zum Verständnis sozialer Kooperation bieten.

## Literatur

Breyer, Thiemo (2013): „Empathie und ihre Grenzen: Diskursive Vielfalt – phänomenale Einheit?“ In: ders. (Hg.): *Grenzen der Empathie. Philosophische, psychologische und anthropologische Perspektiven*. München: Fink. S. 13-42.

- Bullinger, Anke F./Zimmermann, Felicitas/Kaminski, Juliane/Tomasello, Michael (2011): „Different social motives in the gestural communication of chimpanzees and human children“. In: *Developmental Science* (14)1, S. 58–68.
- Chomsky, Noam (1965): *Aspects of the theory of syntax*. Cambridge: MIT Press.
- De Jaeger, H./Di Paolo, E. (2007): „Participatory sense-making. An enactive approach to social cognition“. In: *Phenomenology and the Cognitive Sciences* 6, S. 485–507.
- Dotov, Dobromir G./Nie, Lin/Chemero, Anthony (2010): „A Demonstration of the Transition from Ready-to-Hand to Unready-to-Hand“. In: *PLoS ONE* (5)3, e9433.
- Dreyfus, Hubert (1993): „Heidegger’s Critique of the Husserl/Searle Account of Intentionality“. In: *Social Research* (60)1, S. 17-38.
- Froese, T./Fuchs, T. (2012): „The extended body: A case study in the neurophenomenology of social interaction“. In: *Phenomenology and the Cognitive Sciences* 11, S. 205-235.
- Fuchs, T./De Jaeger, H. (2009): „Enactive intersubjectivity: Participatory sense-making and mutual incorporation“. In: *Phenomenology and the Cognitive Sciences* 8, S. 465-486.
- Garfinkel, Harold (1967): *Studies in ethnomethodology*. Englewood Cliffs: Prentice-Hall.
- Garfinkel, Harold (1996): „Ethnomethodology’s Program“. In: *Social Psychology Quarterly* (59)1, S. 5-21.
- Garfinkel, Harold (2002): *Ethnomethodology’s program: Working out Durkheim’s aphorism*. Lanham: Rowman & Littlefield.
- Gibbs Jr., Raymond W. (2001): „Intentions as Emergent Products of Social Interactions“. In: *Intentions and Intentionality: Foundations of Social Cognition*, hg. v. Malle, Bertram F./Moses, Louis J./Baldwin, Dare A. Cambridge: MIT Press. S. 105-122.
- Girke, Felix/Meyer, Christian (2011): „Introduction“. In: Meyer, Christian/ Girke, Felix (Hg.): *The Rhetorical Emergence of Culture*. Oxford: Berghahn, S. 1-38.
- Goffman, Erving (1959): *The Presentation of self in everyday life*. New York: Doubleday Anchor.
- Goffman, Erving (1974): *Frame Analysis: An Essay on the Organization of Experience*. New York: Harper and Row.
- Goffman, Erving (1983): „The Interaction Order.“ In: *American Sociological Review* 48, 1, S. 1-17.
- Heritage, John (2008): „Conversation Analysis as Social Theory“. In: Turner, B. (Hg.): *The New Blackwell Companion to Social Theory*. Oxford: Blackwell, S. 300-320.
- Hume, David [1739] 1992. *A treatise on human nature*. Oxford: Clarendon.
- Hutchins, Edwin L. (1995). *Cognition in the Wild*. Cambridge: MIT Press.
- Jensen, K./Hare, B./Call, J./Tomasello, M. (2006): „What’s in it for me? Self-regard precludes altruism and spite in chimpanzees“. In: *Proc. R. Soc. B*, (273)1589, S. 1013-1021.
- Joas, Hans (1996): *Die Kreativität des Handelns*. Frankfurt/M.: Suhrkamp.
- Levinson, Stephen C. (1995): „Interactional biases in human thinking“. In: Goody, E.N. (Hg.): *Social intelligence and interaction*. Cambridge: Cambridge University Press, S. 221-226.
- Loenhoff, Jens (2010): „Fundierende Ebenen der Koorientierung und der Handlungskoordination“. In: Böhle, Fritz/Wehrich, Margit (Hg.): *Die Körperlichkeit sozialen Handelns. Soziale Ordnung jenseits von Normen und Institutionen*. Bielefeld: Transcript. S. 59-78.
- Loenhoff, Jens (2012): „Der Körper als Generator vorreflexiver Gewissheit und Medium der Sinnkonstitution“. In: Renn, Joachim/Ernst, Christoph/Isenböck, Peter (Hg.): *Konstruktion und Geltung: Beiträge zu einer postkonstruktivistischen Sozial- und Medientheorie*. Wiesbaden: VS Verlag, S. 183-203.

- Melis, A.P./ Hare, B./Tomasello, M. (2006): „Chimpanzees Recruit the Best Collaborators“. In: *Science* 311, S. 1297-1300.
- Merleau-Ponty, Maurice (2007): „Der Philosoph und sein Schatten“. In: ders.: *Zeichen*. Hamburg: Meiner, S. 233-264, (frz. Original 1960).
- Meyer, Christian (2014): „Metaphysik der Anwesenheit.‘ Zur Universalitätsfähigkeit soziologischer Interaktionsbegriffe“. In: *Zeitschrift für Soziologie*. Sonderheft „Interaktion-Organisation-Gesellschaft“, S. 321-345. Hg. v. Heintz, B./Tyrell, H.
- Meyer, Christian/v. Wedelstaedt, Ulrich (2013): „Skopische Sozialität: Sichtbarkeitsregime und visuelle Praktiken im Boxen“. In: *Soziale Welt* (64)1/2, S. 69-95, Themenheft „Visuelle Soziologie“, hg. v. Baer A./Schnettler; B.
- Meyer, Christian/v. Wedelstaedt, Ulrich (2015): „Teamsubjekte: Körperlich-rituelle Mechanismen der Vergemeinschaftung im Spitzensport“. In: Gugutzer, R./Staaack, M. (Hg.): *Körper und Ritual: Sozial- und kulturwissenschaftliche Zugänge und Analysen*. Wiesbaden: Springer VS. S. 97-124.
- Mills, C. Wright (1940): „Situating Actions and Vocabularies of Motive“. In: *American Sociological Review*, S. 904 - 913.
- Rawls, Anne Warfield (2005): „Garfinkel’s Conception of Time“. In: *Time & Society* (14)2-3, S. 163-190.
- Renn, Joachim (2006): „Rekonstruktion statt Repräsentation. Der ‚pragmatische Realismus‘ John Deweys und die gesellschaftstheoretische Revision des wissenssoziologischen Konstruktivismus“. In: *Soziologische Revue* (28)6, S. 13-37.
- Schegloff, Emanuel A. (1992): „Repair after Next Turn: The Last Structurally Provided Defense of Intersubjectivity in Conversation“. In: *American Journal of Sociology* (97)5, S. 1295-1345.
- Schmid, Hans Bernhard/Schweikard, David P. (Hg.) (2009): *Kollektive Intentionalität*. Frankfurt/M.: Suhrkamp.
- Schütz, Alfred (1974): *Der sinnhafte Aufbau der sozialen Welt. Eine Einleitung in die verstehende Soziologie*. Frankfurt/M.: Suhrkamp, (orig. 1932).
- Searle, John (1983): *Intentionality: An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Shotter, John (2011): „The Dance of Rhetoric Dialogic Selves and Spontaneously Responsive Expressions“. In: Meyer, Christian/ Girke, Felix (Hg.): *The Rhetorical Emergence of Culture*. Oxford: Berghahn. S. 35-51.
- Slaby, Jan (2014): „Empathy’s blind spot“. In: *Medicine, Healthcare and Philosophy* (17)2, S. 249-58.
- Tomasello, M./Rakoscy, H. (2003): „What Makes Human Cognition Unique? From Individual to Shared to Collective Intentionality“. In: *Mind & Language* 18, S. 121-147.
- Tomasello, M. (1996): „Do apes ape?“. In: Galef, J./Heyes, C. (Hg.): *Social Learning in Animals: The Roots of Culture*. Academic Press.
- Tomasello, M. (2008): *Origins of Human Communication*. MIT Press.
- Tomasello, M. (2009): *Why We Cooperate*. MIT Press
- Tomasello, M. (2014): *A Natural History of Human Thinking*. Harvard UP.
- Tomasello, Michael (2006): „Why don’t apes point?“. In: Enfield, N.J./Levinson, S.C. (Hg.): *Roots of Human Sociality: Culture, cognition and interaction*. Oxford & New York: Berg, S. 506-524.
- Tomasello, Michael/Carpenter, Malinda (2007): „Shared intentionality“. In: *Developmental Science* (10)1, S. 121-125.

- Tomasello, Michael/Carpenter, Malinda/Call, Josep/Behne, Zanya/Moll, Henrike (2005): „Understanding and sharing intentions: The origins of cultural cognition“. In: *Behavioral and Brain Sciences* 28, S. 675 - 691.
- Tuomela, Raimo/Miller, Kaarlo (1988): „We-Intentions“. In: *Philosophical Studies* 53, S. 367 - 389.
- Warneken, F./Tomasello, M. (2006): „Altruistic Helping in Human Infants and Young Chimpanzees“. In: *Science* 311, S. 1301 – 1303.
- Warneken, F./Hare, B./Melis, A.P./Hanus, D./Tomasello, M. (2007): „Spontaneous altruism by chimpanzees and young children“. In: *PLoS Biology* (5)7, S. 1414 – 1420.
- Wolff, Stephan (1976): *Der rhetorische Charakter sozialer Ordnung: Selbstverständlichkeit als soziales Problem*. Berlin. Duncker & Humblot

## Anhang: Verwendete Transkriptionszeichen, Lemmata und Siglen

(.)	Mikropause
(-); (--); (---)	Pausen von ca. 0,25; 0,5; 0,75 Sek. Länge (geschätzt)
(1,5)	Pause in gemessener Länge
=	Unmittelbarer Anschluss
:	Vokallängung
beTOnung	Betonte Silben in Großschrift
((klatscht))	Paraverbale Äußerungen und Ereignisse
( )	Unverständliche Passage bzw. unsichere Transkription
(selber)	Unsichere Transkription
((...))	Auslassung im Transkript
Pfeil am Kopf	Blickrichtung
Pfeil am Körper	Bewegungsrichtung eines Körpers