

**Determinanten sozialer Entscheidungsprozesse – der
Einfluss von sozio-kontextuellen Faktoren und
Persönlichkeitseigenschaften auf altruistisches
Verhalten, Fairness und Kooperation in
experimentellen Spielparadigmen**

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Jonathan Frieder Scheeff, M.Sc.
aus Leonberg

Tübingen
2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation:

18.10.2019

Dekan:

Prof. Dr. Wolfgang Rosenstiel

1. Berichterstatter:

Prof. Dr. Michael Schönenberg

2. Berichterstatter:

Prof. Dr. Martin Hautzinger

Inhaltsverzeichnis

1. Einleitung.....	8
1.1 Paradigmen zur Untersuchung sozialer Entscheidungsprozesse.....	8
1.1.1 Dictator Game	9
1.1.2 Ultimatum Game	10
1.1.3 Gefangenendilemma.....	11
1.1.4 Give Some Dilemma	12
1.2 Determinanten sozialer Entscheidungsprozesse.....	13
1.2.1 Sozio-kontextuelle Faktoren.....	13
1.2.1.1 Gruppenzugehörigkeit.....	14
1.2.1.2 Reziprozität.....	14
1.2.2 Persönlichkeitseigenschaften.....	14
1.2.2.1 Empathie.....	15
1.2.2.2 Dissozialität.....	15
1.3 Zielsetzung	19
2. Ergebnisse.....	20
2.1 Studie 1: „Empathy promotes altruistic behavior in economic interactions“	20
2.2 Studie 2: „Investigating social-contextual determinants of cooperation in incarcerated violent offenders“	21
2.3 Studie 3: „Effects of group affiliation on neural signatures of fairness norm violations in antisocial violent offenders“	23
3. Diskussion	25
3.1 Determinanten sozialer Entscheidungsfindung	25
3.1.1 Empathie.....	26
3.1.2 Reziprozität	28
3.1.3 Gruppenzugehörigkeit.....	29
3.1.4 Affektive Korrelate.....	31
3.2 Soziale Entscheidungsfindung bei antisozialen Individuen	32
3.2.1 Abweichendes Verhalten durch strategisches und rationales Handeln	33
3.2.2 Einfluss sozio-kontextueller Determinanten	36
3.2.3 Neurophysiologische Aspekte	37
3.3 Einfluss von Emotionen und Empathie bei antisozialen Individuen	39

3.4 Kritische Würdigung der verwendeten Paradigmen	41
3.5 Fazit	44
Literatur	45
Appendix.....	53

Zusammenfassung

Im Rahmen dieser Dissertation wurden in drei Studien verschiedene Aspekte sozialer Entscheidungsprozesse vor dem Hintergrund unterschiedlicher sozio-kontextueller und individueller Charakteristika erforscht. Studie 1 untersuchte bei gesunden Personen, inwiefern sich das Auslösen von Empathie auf altruistisches Verhalten auswirkt. Dabei konnte erstmals nachgewiesen werden, dass das psychologische Konzept der Empathie auch in ökonomischen Interaktionen einen Einfluss haben kann. In Studie 2 wurden gesunde Individuen und inhaftierte Gewaltstraftäter verglichen. Dabei war es von Interesse inwiefern sich das kooperative Verhalten dieser beiden Populationen durch sozio-kontextuelle Faktoren in Form von Gruppenzugehörigkeit und Reziprozität beeinflussen ließen. Studie 3 untersuchte, ob gesunde Personen und inhaftierte Gewaltstraftäter unterschiedlich auf Fairnessnormverletzungen in einem Gruppenkontext reagieren. Es konnte in Studie 2 und 3 ein erheblicher Einfluss dieser sozio-kontextueller und individueller Charakteristika auf die soziale Entscheidungsfindung nachgewiesen werden, wobei hervorzuheben ist, dass sowohl gesunde wie auch dissoziale Individuen gleichermaßen von diesen Faktoren beeinflusst wurden. Die referierten Befunde werden im Rahmen dieser Arbeit in einer Gesamtdiskussion in die bestehende Literatur eingeordnet und diskutiert.

Summary

In the scope of this work, three studies investigated different aspects of social decision-making considering various socio-contextual and individual characteristics. Study 1 recruited healthy individuals and asked whether an empathy induction would promote altruism. We were the first to demonstrate, that the psychological concept of empathy had a substantial influence in economic interactions. Study 2 compared healthy individuals with incarcerated violent offenders regarding a potential modulation of cooperative behavior through the socio-contextual factors group affiliation and reciprocity. Study 3 investigated how healthy individuals and incarcerated violent offenders react to fairness norm violations in a group context. There was a considerable influence of socio-contextual and individual characteristics on social decision-making. Yet, it is important to note, that healthy and dissocial individuals were affected in the same manner. The reported results are discussed by reference to the current literature in a general discussion.

Liste der Publikationen der Dissertation

- Klimecki, O. M., Mayer, S. V., Jusyte, A., Scheeff, J., & Schönberg, M. (2016). Empathy promotes altruistic behavior in economic interactions. *Scientific reports*, 6, 31961. doi:10.1038/srep31961
- Scheeff, J., Jusyte, A., & Schönberg, M. (2018). Investigating social-contextual determinants of cooperation in incarcerated violent offenders. *Scientific reports*, 8(1), 17204. . doi:10.1038/s41598-018-35450-z
- Scheeff, J., Rauss, K., & Schönberg, M. (submitted). Effects of group affiliation on neural signatures of fairness norm violations in antisocial violent offenders.

1. Einleitung

Der menschliche Alltag besteht aus einer Vielzahl von sozialen Entscheidungen, die ein Individuum fällen muss. Dabei kann es sich um triviale Dinge handeln, wie beispielsweise den Nachbarn zu grüßen oder einem Familienmitglied bei einem Umzug zu helfen. Es gibt jedoch auch weitreichendere Beschlüsse wie Geld an eine Wohltätigkeitsorganisation zu spenden oder sich für den Umweltschutz zu engagieren. Anhand dieser Beispiele kann bereits festgestellt werden, dass Entscheidungen durch eine Vielzahl von Faktoren beeinflusst werden. Hier scheint es eine Rolle zu spielen, ob man den Interaktionspartner ein einziges Mal oder wiederholt sieht und ob er einem nahesteht. Zudem scheint es intuitiv, dass verschiedene Gefühle eine zentrale Determinante des Verhaltens sind, zum Beispiel Mitgefühl und Empathie bei Spenden.

Verschiedene wissenschaftliche Disziplinen beschäftigen sich intensiv mit der Frage, wie diese Entscheidungsprozesse funktionieren, dabei wurde in der Ökonomie erstaunlich lange die Rolle von Emotionen und kognitiven Prozessen kaum beachtet. Lange galt der Mensch als so genannter Homo oeconomicus („Wirtschaftsmensch“) oder auch als rationaler Agent, der stets den eigenen Nutzen maximiert und weitestgehend egoistisch handelt (Persky, 1995). Emotionen und andere Faktoren spielen in diesem theoretischen Modell keine wesentliche Rolle. Diese Sichtweise wurde in den letzten Jahrzehnten zunehmend in Frage gestellt und durch eine Vielzahl an Experimenten widerlegt. Diese Studien konnten zeigen, dass Menschen sich durch kognitive Verzerrungen (so genannter *bias*, zu Deutsch Verzerrung oder Verzerrungseffekt) sowie andere psychologische Faktoren sowie Emotionen und eben nicht durch pure Vernunft leiten lassen (Balliet, Wu, & De Dreu, 2014; Eisenberg & Miller, 1987; Gintis, 2000).

1.1 Paradigmen zur Untersuchung sozialer Entscheidungsprozesse

Um Entscheidungsprozesse in sozialen Situationen zu untersuchen, wurden verschiedenste Paradigmen entwickelt und getestet. Bereits in den 80er Jahren wurden die ersten dieser Spiele entworfen und mittlerweile tausendfach in Studien angewandt. Gibt man beispielsweise in der Suchmaschine Google Scholar die unten genannten vier Spiele ein, so finden sich über 80000 Einträge zu Studien mit diesen Schlagwörtern (Stand Juli 2019). Diese gewaltige Anzahl wissenschaftlicher Untersuchungen

unterschiedlicher Disziplinen belegt das große Interesse sowie die Bedeutung dieser Methoden. Im Folgenden soll auf die prominentesten sowie die im Rahmen dieser Dissertation relevanten Paradigmen eingegangen werden. Allen vorgestellten Paradigmen ist gemeinsam, dass jeweils zwei Personen miteinander interagieren beziehungsweise spielen.

1.1.1 Dictator Game

In diesem Paradigma kann ein Spieler (der so genannte Diktator) nach Belieben einen bestimmten Geldbetrag oder anderweitig zur Verfügung stehende Ressourcen zwischen sich und dem anderen Spieler aufteilen, welcher dieses Angebot nur annehmen kann (Forsythe, Horowitz, Savin, & Sefton, 1994). Dieses Paradigma soll altruistisches Verhalten messen. Ein rationaler Spieler im Sinne des Homo oeconomicus würde seinem Gegenpart nichts oder nur sehr wenig abgeben, um selbst möglichst viel zu erhalten. Dennoch konnte in einer Meta-Analyse mit hunderten ausgewerteten Studien festgestellt werden, dass gesunde Personen im Durchschnitt 30% der Ressourcen abgeben (Engel, 2011).

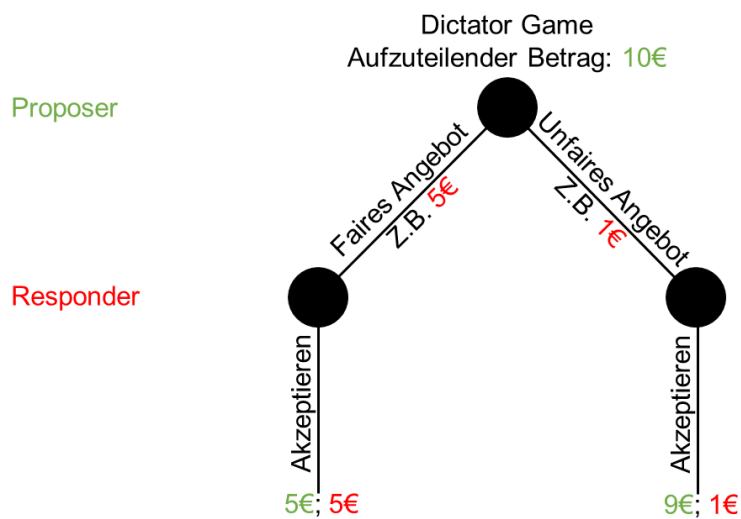


Abbildung 1. Schematische Darstellung des Verlaufs sowie beispielhafter Ergebnisse im Dictator Game.

1.1.2 Ultimatum Game

Ein weiteres prominentes Paradigma ist das Ultimatum Game. Hier schlägt ein Spieler (der so genannte *proposer*, zu Deutsch Vorschlagender) ebenfalls nach Belieben eine Aufteilung der zur Verfügung stehenden Ressourcen vor, die der Mitspieler (der so genannte *responder*, zu Deutsch der Beantwortende oder Empfänger) annehmen oder ablehnen kann (Güth, Schmittberger, & Schwarze, 1982). Nimmt der Responder das Angebot an, wird der Betrag wie vorgeschlagen aufgeteilt. Lehnt er ab, bekommt keiner der beiden Spieler etwas. Dieses Paradigma soll daher die Fähigkeit zum Verständnis von Fairnessnormen sowie die Reaktion auf deren Verletzungen messen. Auch hier würde ein rationaler Proposer so wenig Punkte wie möglich abgeben und ein rationaler Responder selbst extrem niedrige Angebote annehmen, da wenig immer noch besser als nichts ist. Auch in diesem Spiel widersprechen typische Befunde diesen Annahmen. Die meisten Responder lehnen Angebote unter 30% des Gesamtbetrags häufig ab und zudem geben Proposer im Durchschnitt mit 50% einen wesentlichen Teil der Ressourcen ab (Camerer, 2003; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003; Yamagishi et al., 2009).

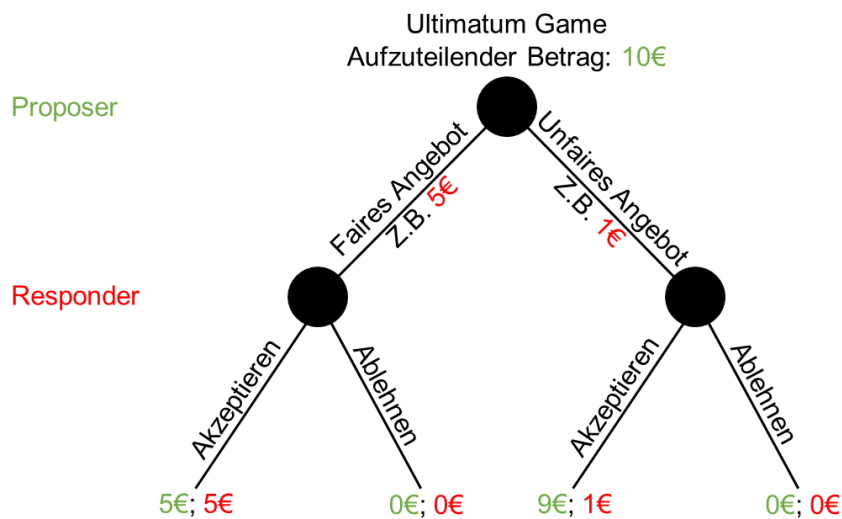


Abbildung 2. Schematische Darstellung des Verlaufs sowie beispielhafter Ergebnisse im Ultimatum Game.

Die passive Rolle des Responders in diesem Spiel erlaubt es zudem ereigniskorrelierte Potentiale (EKP) im Elektroenzephalogramm (EEG) auf verschiedene Angebote zu erfassen. In der Literatur wurde hier häufig die so genannte Medial Frontal Negativity (MFN) untersucht, welche eine negative Deflektion im fronto-zentralen Bereich darstellt und meist 250-350ms nach nachteiligen im Vergleich zu vorteilhaften Angeboten auftritt (Polezzi et al., 2008). Bisher wird angenommen, dass die MFN eine affektive Evaluation negativer Ereignisse (Boksem, Tops, Wester, Meijman, & Lorist, 2006; Gehring & Willoughby, 2002; Masaki, Takeuchi, Gehring, Takasawa, & Yamazaki, 2006) oder die Verletzung sozialer Normen und Erwartungen darstellt (Boksem & De Cremer, 2010; Pfabigan, Alexopoulos, Bauer, & Sailer, 2011; Potts, Martin, Burton, & Montague, 2006; Sato et al., 2005). Im Ultimatum Game wird dieses EKP daher typischerweise nach unfairen im Vergleich zu fairen Angeboten beobachtet (Boksem & De Cremer, 2010; Hewig et al., 2011).

1.1.3 Gefangenendilemma

Das Gefangenendilemma (englisch Prisoners Dilemma) ist ein weiteres Paradigma aus der Spieltheorie. Es modelliert eine Situation, bei der zwei Verbrecher gefangen wurden und nun entweder gestehen oder schweigen können ohne miteinander kommunizieren zu können. Gestehen beide, erhalten sie jeweils eine Strafe von zwei Jahren Gefängnis. Schweigen beide, erhalten sie eine minimale Strafe von 6 Monaten, da es nicht genug Beweise gibt. Gesteht jedoch nur einer und der andere schweigt, wird der Kronzeuge freigelassen, während der Schweigende durch die Zeugenaussage des anderen zu einer Maximalstrafe von fünf Jahren verurteilt wird (siehe Rapoport, 1989). Dieses Spiel existiert in verschiedensten Varianten, zentral ist jedoch immer, dass zwei Spieler miteinander kooperieren oder sich verraten können ohne miteinander sprechen zu können (Axelrod & Hamilton, 1981; Rapoport & Chammah, 1965). Dieses Dilemma zwischen Eigeninteresse und kollektivem Ergebnis wird verwendet, um Kooperation zwischen Individuen zu messen. Dabei liegen die Kooperationsraten um 50% (Sally, 1995), diese Zahlen variieren erheblich je nach verwendeter Payoff-Matrix. Auch hier sind also individuell rationale Entscheidungen häufig mit einem kollektiv schlechten Ergebnis verbunden (Kollock, 1998). Ein Nachteil dieses Paradigmas besteht in der

kategorialen Erfassung von Kooperation im Sinne von „kooperieren“ oder „nicht kooperieren“.

		Player B	
		Kooperieren	Verraten
Player A	Kooperieren	7€ / 7€	0€ / 10€
	Verraten	10€ / 0€	2€ / 2€

Abbildung 3. Schematische Darstellung einer beispielhaften Gewinnmatrix des Gefangenendilemmas.

1.1.4 Give Some Dilemma

Der Nachteil einer kategorialen Erfassung von Kooperation wird im so genannten Give Some Dilemma (zu Deutsch Abgabe Dilemma) verbessert umgesetzt: Hier müssen zwei Spieler jeweils zwischen sich und dem anderen zehn Geldeinheiten aufteilen (wie im Dictator Game oder Ultimatum Game), wobei jeder geteilte Punkt für den jeweiligen Mitspieler doppelt zählt. Daher wird hier als die Anzahl an geteilten Punkten als kontinuierliches beziehungsweise numerisches Maß für Kooperation herangezogen (Van Lange & Kuhlman, 1994; Van Lange, Ouwerkerk, & Tazelaar, 2002). In diesem sozialen Dilemma hat ein Individuum einen größeren Outcome, wenn er oder sie weniger Punkte abgibt (individuelle Rationalität). Auf der anderen Seite haben beide Spieler mehr davon, wenn jeder dem anderen mehr Geldeinheiten abgibt (kollektive Rationalität) (Van Lange et al., 2002). In anderen Worten gibt es in diesem komplexen Szenario also nicht die eine rationale Lösung. Dies spiegelt sich auch in den Befunden mit diesem Paradigma wieder: Je nach Persönlichkeitseigenschaften der Spieler und Spielinstruktionen geben Individuen in einem typischen Give Some Dilemma zwischen

drei und fast acht Geldeinheiten ab (Van Lange, Klapwijk, & Van Munster, 2011; Van Lange et al., 2002).

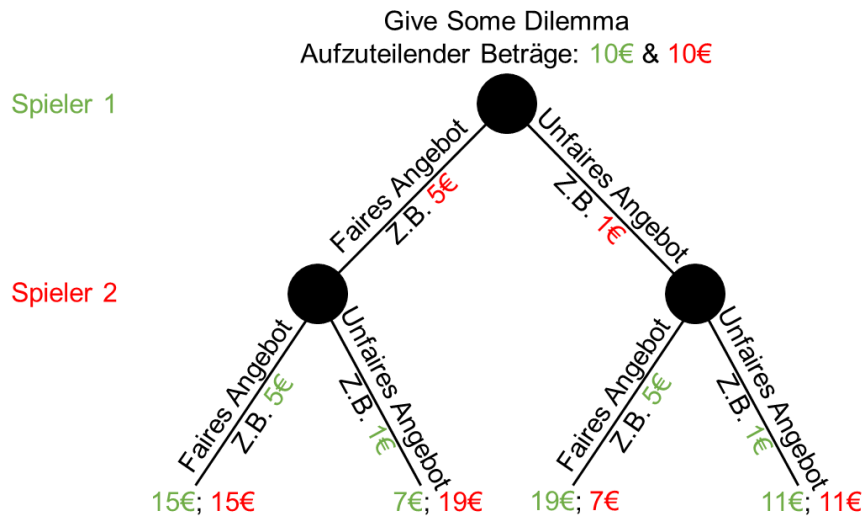


Abbildung 4. Schematische Darstellung des Verlaufs sowie beispielhafter Ergebnisse im Give Some Dilemma.

1.2 Determinanten sozialer Entscheidungsprozesse

Zusammenfassend kann festgestellt werden, dass Menschen sich in spieltheoretischen Paradigmen bei weitem großzügiger, fairer und kooperativer verhalten als erwartet. Eine spannende Forschungsfrage ist nun, welche Determinanten den Entscheidungen für solche Verhaltensweisen zugrunde liegen. Dabei kann zwischen Persönlichkeitseigenschaften der Spieler und sozio-kontextuellen Faktoren unterschieden werden.

1.2.1 Sozio-kontextuelle Faktoren

Die bisherige Forschung konnte eine Reihe von sozio-kontextuellen Faktoren identifizieren, die das Verhalten in sozialen Entscheidungssituationen beeinflussen. Beispielsweise konnte in einer Meta-Analyse zum Dictator Game gezeigt werden, dass es eine Rolle spielt, ob der Empfänger notleidend ist, die Entscheidung in einer Gruppe gefällt wird oder der Diktator identifizierbar ist (Engel, 2011). Es scheint außerdem

wichtig zu sein, mit wem man es zu tun hat und wie oft man dieser Person begegnet. Diese Charakteristika einer sozialen Interaktion, nämlich die Gruppenzugehörigkeit des Interaktionspartners und die Anzahl an erwarteten Interaktionen mit diesem wurden im Rahmen dieser Dissertation untersucht. Beide Variablen stellen etablierte psychologische Faktoren dar, welche das Verhalten von Menschen beeinflussen und zum Beispiel kooperatives Verhalten fördern können (Balliet et al., 2014; Wilson, 1971).

1.2.1.1 Gruppenzugehörigkeit

Es konnte gezeigt werden, dass Mitglieder der eigenen Gruppe denen einer anderen Gruppe gegenüber bevorzugt werden (so genannter Eigengruppen-Bias oder Eigengruppen-Bevorzugung) (Brewer, 1979). Dies geschieht vor allem aus dem Wunsch, positive Verhältnisse in der eigenen Gruppe aufrechtzuhalten und zu stärken (Brewer, 1999). Diese Unterscheidung zwischen der eigenen und fremden Gruppe scheint zudem recht basal zu sein: Sie tritt bereits in minimalistischen Settings auf, bei denen die Teilnehmer zufällig zu Gruppen zugeteilt werden, nicht miteinander kommunizieren können oder keine anderen Mitglieder der eigenen oder fremden Gruppe kennen (Brewer, 1979, 1999; Tajfel, Billig, Bundy, & Flament, 1971; Tajfel & Turner, 1979).

1.2.1.2 Reziprozität

Des Weiteren führt bereits die bloße Antizipation einer wiederholten Interaktion zu kooperativerem Verhalten im Vergleich zu einer einfachen Interaktion (Van Lange et al., 2011). Diese Bereitschaft zu erhöhter Kooperation in wiederholten Interaktionen erklärt sich durch die Vorteile der Kooperation sowie Möglichkeit zur Bestrafung von unkooperativen Verhaltensweisen, das heißt die sich hieraus ergebende Reziprozität (oder auch Wechselseitigkeit) begünstigt prosoziales Verhalten (Axelrod & Hamilton, 1981). Darüber hinaus zeigen Individuen einen noch stärkeren Eigengruppen-Bias, wenn das eigene Ergebnis im Sinne der Reziprozität vom Verhalten des Mitspielers abhängt (Balliet et al., 2014).

1.2.2 Persönlichkeitseigenschaften

Innerhalb einer Person können verschiedenste Charakteristika mit sozialer Entscheidungsfindung assoziiert werden. Aus unterschiedlichen Paradigmen konnten

zahlreiche Persönlichkeitseigenschaften identifiziert werden, die mit dem Verhalten in ökonomischen Spielen assoziiert sein könnten. Als Beispiele können hier folgende Befunde genannt werden: Im Ultimatum Game wurden egoistische Persönlichkeitseigenschaften mit Forderungen nach höheren Angeboten in Verbindung gebracht (Brandstätter & Königstein, 2001), prosoziale beziehungsweise individualistische Orientierungen wirkten sich auf die Kooperationsbereitschaft im Give Some Dilemma aus (Van Lange et al., 2011) und Ehrlichkeit sowie Bescheidenheit sagten altruistisches Verhalten in einem Dictator Game vorher (Hilbig, Thielmann, Hepp, Klein, & Zettler, 2015). Da eine bloße Aufzählung dieser umfangreichen Persönlichkeitseigenschaften den Rahmen dieser Dissertation überschreiten würde, wird im Folgenden auf zwei zentrale Eigenschaften, der Empathiefähigkeit und Dissozialität, eingegangen.

1.2.2.1 Empathie

Mit empathischen Persönlichkeitseigenschaften ist vor allem die Fähigkeit gemeint, Gefühle anderer nachzuempfinden (im Sinne der emotionalen Empathie), was wiederum mit altruistischem Verhalten in Verbindung gebracht wurde (de Waal, 2008; Eisenberg & Fabes, 1990). Diese Annahme wurde in einer Metaanalyse bestätigt indem gezeigt werden konnte, dass empathische Zustände und (wenn auch zu einem geringeren Ausmaß) empathische Eigenschaften altruistisches Verhalten vorhersagen (Eisenberg & Miller, 1987). Dennoch gibt es bisher kaum Untersuchungen, welche die psychologischen Konzepte von Empathie mit den ökonomischen Maßen für Prosozialität (beispielsweise im Sinne des im Dictator Game gemessenen Altruismus) verbinden.

1.2.2.2 Dissozialität

Menschen mit dissozialen Persönlichkeitseigenschaften fehlt oftmals die Fähigkeit zur Empathie (Jolliffe & Farrington, 2004). Damit sind Individuen gemeint, die sich hauptsächlich für eigene Interessen einsetzen, ohne Rücksicht auf Bedürfnisse oder Gefühle anderer. In der Literatur wird dabei häufig von einer so genannten dunklen Triade aus den Persönlichkeitseigenschaften Narzissmus, Psychopathie und Machiavellismus gesprochen (Paulhus & Williams, 2002). Diese Eigenschaften sind mittlerweile häufig in der Allgemeinbevölkerung untersucht und diskutiert worden (Furnham, Richards, & Paulhus, 2013). Unter anderem zeigen sich hier

Zusammenhänge zwischen Defiziten bezüglich empathischer Fähigkeiten und der dunklen Triade (Wai & Tiliopoulos, 2012). In dieser Dissertation wird vor allem auf eine ausgeprägte Form dieser Eigenschaften im Sinne von klinisch relevanten dissozialen Persönlichkeitseigenschaften eingegangen. Diese Form der Psychopathologie wird vor allem im Rahmen der klinischen und forensischen Psychologie als so genannte antisoziale Persönlichkeitsstörung erforscht und aufgrund der Schwere dieser Störung sowie dem bisherigen fehlen zuverlässiger Therapiemöglichkeiten (Gibbon et al., 2010) scheint es hier umso dringlicher, grundlagenwissenschaftliche Fragestellungen zu untersuchen.

Individuen mit antisozialer Persönlichkeitsstörung zeichnen sich durch ein tiefgreifendes Muster von Missachtung und Verletzungen der Rechte anderer aus, das seit dem 15. Lebensjahr auftritt (American Psychiatric Association, 2013). Zudem müssen mindestens drei der folgenden Kriterien erfüllt sein: (1) Versagen, sich in Bezug auf gesetzmäßiges Verhalten gesellschaftlichen Normen anzupassen, was sich in wiederholtem Begehen von Handlungen äußert, die einen Grund für eine Festnahme darstellen. (2) Falschheit, die sich in wiederholtem Lügen, dem Gebrauch von Decknamen oder dem Betrügen anderer zum persönlichen Vorteil oder Vergnügen äußert. (3) Impulsivität oder Versagen, vorausschauend zu planen. (4) Reizbarkeit und Aggressivität, die sich in wiederholten Schlägereien oder Überfällen äußert. (5) Rücksichtslose Missachtung der eigenen Sicherheit bzw. der Sicherheit anderer. (6) Durchgängige Verantwortungslosigkeit, die sich im wiederholten Versagen zeigt, eine dauerhafte Tätigkeit auszuüben oder finanziellen Verpflichtungen nachzukommen. (7) Fehlende Reue, die sich in Gleichgültigkeit oder Rationalisierungen äußert, wenn die Person andere Menschen gekränkt, misshandelt oder bestohlen hat. Für die Diagnose der antisozialen Persönlichkeitsstörung nach dem DSM V muss die betreffende Person mindestens 18 Jahre alt und zudem eine Störung des Sozialverhaltens bereits vor Vollendung des 15. Lebensjahres erkennbar gewesen sein. Zusätzlich sollte das gezeigte antisoziale Verhalten nicht ausschließlich im Verlauf einer Schizophrenie oder einer bipolaren Störung auftreten.

Die antisoziale Persönlichkeitsstörung stellt mit einer Prävalenz in der Allgemeinbevölkerung (je nach Studie) zwischen 1% und 6% (Coid, Yang, Tyrer,

Roberts, & Ullrich, 2006; Robins & Price, 1991; Torgersen, Kringlen, & Cramer, 2001) eine häufige Störung dar. Männer sind im Vergleich zu Frauen häufiger betroffen, zudem findet sich eine massive Häufung unter Gefängnisinsassen (Compton, Conway, Stinson, Colliver, & Grant, 2005; Fazel & Danesh, 2002). Die antisoziale Persönlichkeitsstörung scheint mit erheblichen Komorbiditäten, vor allem mit anderen Persönlichkeitsstörungen, Substanzmissbrauch (Compton et al., 2005) sowie anderen Achse 1 Störungen assoziiert zu sein (Goodwin & Hamilton, 2003; Lenzenweger, Lane, Loranger, & Kessler, 2007). Zudem steht diese Erkrankung im Zusammenhang mit einer erhöhten Sterblichkeitsrate, gerade im jüngeren Lebensalter und vor allem durch rücksichtsloses Verhalten (Black, Baumgard, Bell, & Kao, 1996). Des Weiteren stellt das Konstrukt der Psychopathie eine noch stärker beeinträchtigte Subgruppe innerhalb des antisozialen Spektrums dar (Coid & Ullrich, 2010). Charakteristisch sind spezifische Eigenschaften im interpersonell-affektiven Bereich (manipulatives Verhalten, Gefühlskälte, Mangel an Empathie) sowie antisozial-deviante Auffälligkeiten (unzureichende Verhaltenskontrolle, Verantwortungslosigkeit). Die Häufigkeit von Psychopathen in der Allgemeinbevölkerung wird mit einer Prävalenz von 0.6 als recht niedrig geschätzt (Coid, Yang, Ullrich, Roberts, & Hare, 2009), während in Gefängnissen von Prävalenzraten zwischen 15 und 25% ausgegangen wird (Hare, 2003). Trotz massiver individueller sowie gesellschaftlicher Kosten, die als bis zu zehnmal höher für antisoziale Individuen im Vergleich zu Gesunden beziffert werden (Scott, Knapp, Henderson, & Maughan, 2001), ist die momentane Versorgungslage als unzureichend einzuschätzen, da keine wirksame Behandlung für die antisoziale Persönlichkeitsstörung vorliegt (Gibbon et al., 2010).

Defizite in der sozialen Informationsverarbeitung wurden in der Vergangenheit extensiv untersucht um aggressives und gefühlloses Verhalten in diesen Populationen zu erklären (Dawel, O’Kearney, McKone, & Palermo, 2012; Hoppenbrouwers, Bulten, & Brazil, 2016; Marsh & Blair, 2008). Dennoch ist es überraschend, dass bisher kaum Untersuchungen zu verhaltensrelevanten Maßen wie Kooperation oder sozialer Regelverletzung vorliegen. Vor dem Hintergrund der einschlägigen Diagnosekriterien wären gerade diese Paradigmen von besonderer Relevanz, um Defizite im Normverständnis und abweichendes Verhalten zu identifizieren. Vor dem Hintergrund

der einschlägigen Diagnosekriterien kann vermutet werden, dass Individuen mit antisozialer Persönlichkeitsstörung Defizite bezüglich empathischer Fähigkeiten aufweisen. Diese Annahme konnte in einer Meta-Analyse bestätigt werden, indem Zusammenhänge zwischen dem Begehen von Straftaten, vor allem Gewaltstraftaten, und Defiziten bezüglich der Empathie festgestellt werden konnten (Jolliffe & Farrington, 2004). Geht man davon aus, dass gerade diese Fähigkeit bei der Entscheidungsfindung in sozialen Situationen eine zentrale Rolle spielt, unterstreicht dieser Befund die Wichtigkeit der Untersuchung von Entscheidungsprozessen bei Menschen mit antisozialer Persönlichkeitsstörung.

1.3 Zielsetzung

Im Rahmen der hier vorgelegten Dissertation sollen die zuvor referierten Persönlichkeitseigenschaften und sozio-kontextuellen Faktoren systematisch untersucht werden. Studie 1 befasst sich mit dem Einfluss von Empathie auf prosoziales Verhalten in Form von Altruismus. Dabei wird das psychologische Konzept der Empathie mit dem ökonomischen Paradigma des Dictator Games in Verbindung gebracht und untersucht, ob der Grad an erlebter Empathie gegenüber einem Rezipienten eine zentrale Bedeutung für prosoziales Verhalten diesem gegenüber hat. Während diese Studie gesunde Individuen untersucht, beschäftigen sich Studie 2 und 3 zusätzlich mit einem Kollektiv, welches für das Gegenteil von Empathie und Altruismus bekannt ist: Individuen mit antisozialen und gewalttätigen Tendenzen. Studie 2 untersucht, in welchem Maß antisoziale Individuen in der Lage sind prosoziales Verhalten zu zeigen und ob sich dieses auf ähnliche Art und Weise wie bei Gesunden durch die Gruppenzugehörigkeit des Mitspielers sowie die Anzahl an erwarteten Interaktionen beeinflussen lässt. Dabei wurde auch der Einfluss von psychopathischen Eigenschaften auf die Entscheidungsfindung explorativ untersucht. In der 3. Studie wird anschließend die Perspektive verändert: Während bisher stets Individuen in der aktiven Rolle des Gebenden betrachtet wurden, wird nun beleuchtet, wie antisoziale Individuen im Vergleich zu Gesunden auf Normverletzungen ihnen gegenüber reagieren. Da neurophysiologische Beeinträchtigungen in der Vergangenheit weitreichend untersucht und belegt wurden, werden hier zusätzlich zu den Verhaltensmaßen elektrophysiologische Prozesse im Gehirn aufgezeichnet, um ein tiefergreifendes Verständnis über zugrundeliegende Mechanismen und Normen zu erhalten. Auch hier wurde der Einfluss von psychopathischen Eigenschaften auf die relevanten Verhaltensmaße und elektrophysiologischen Marker betrachtet.

2. Ergebnisse

2.1 Studie 1: „Empathy promotes altruistic behavior in economic interactions“

In dieser Studie wurde getestet, ob die Empathie-Altruismus Hypothese auch im Bereich der Ökonomie gilt. Dafür wurde ein within-subject Experiment durchgeführt, bei dem das Verhalten in einem typischen einmaligen Dictator Game mit einer neuartigen Version des Dictator Game, dem so genannten empathischen Dictator Game verglichen wurde. In letzterem Paradigma ging den ökonomischen Entscheidungen stets das Auslösen von Empathie (so genannter Empathie-Induktion) gegenüber dem Empfänger oder eine neutrale Kontrollbedingung (ohne Induktion von Empathie) voraus. Hierfür wurde vor jedem Durchgang ein Video mit entsprechendem Inhalt (notleidender Empfänger oder Empfänger in alltäglicher, neutraler Situation) gezeigt. In beiden Versionen wurden die Versuchsteilnehmer aufgefordert Geldeinheiten zwischen sich und den Empfängern aufzuteilen. Dieses Experiment wurde mit einer studentischen Stichprobe durchgeführt. Im Folgenden werden die zentralen Ergebnisse der Studie zusammengefasst.

In dieser Studie konnte zunächst gezeigt werden, dass die experimentelle Manipulation erfolgreich funktionierte. Hierfür wurden die verschiedenen Videos aus dem Empathie Dictator Game und der Kontrollbedingung bezüglich der Auslösung von Empathie sowie negativen und positiven Emotionen verglichen. Wie erwartet lösten die Empathie induzierenden Videos mehr Empathie und negative Gefühle aus, während die Kontrollvideos mehr positive Gefühle bei den Betrachtern hervorriefen. Des Weiteren wurde das Gebeverhalten in den verschiedenen Bedingungen des Dictator Games (Standard one-shot, Empathie induzierend und neutrale Kontrollbedingung) verglichen. Auch hier zeigte sich wie erwartet ein Einfluss der Experimentalbedingung: Während im Standard Dictator Game 30% der Geldeinheiten sowie in der Kontrollbedingung 40% abgegeben wurde, gaben Teilnehmer nach dem Auslösen von Empathie im Durchschnitt über 70% der Geldeinheiten ab. Zusätzlich wurden Regressionsanalysen durchgeführt, um zu überprüfen welche momentanen State (Zustand) das Gebeverhalten beeinflussten. Dabei zeigte sich für die Empathie Bedingung, dass die Empathie gegenüber dem Empfänger das Gebeverhalten vorhersagte (diese Variable erklärte über 30% der Varianz). Negative Gefühle hatten isoliert ebenfalls einen Vorhersagewert. Wurden jedoch beide Variablen in die Analyse miteinbezogen, blieb

nur noch die State Empathie als relevante Determinante übrig. Abschließend hatte die Trait (Eigenschaft) Empathie der Teilnehmer in unserer Studie allerdings keinen Einfluss auf das Gebeverhalten in allen experimentellen Bedingungen.

2.2 Studie 2: „Investigating social-contextual determinants of cooperation in incarcerated violent offenders“

Studie 1 überprüfte die Wichtigkeit von Empathie bei der sozialen Entscheidungsfindung. Geht man von einem Empathiedefizit bei Menschen aus dem antisozialen Spektrum aus (Jolliffe & Farrington, 2004), so kann angenommen werden, dass diese Individuen Auffälligkeiten bei der Entscheidungsfindung aufweisen. Diese Studie hatte daher zum Ziel, kooperatives Verhalten und dessen sozio-kontextuelle Determinanten bei inhaftierten Gewaltstraftätern im Vergleich zu gesunden Kontrollen mittels des Give Some Dilemma zu untersuchen. Zentral für diese Untersuchung war die Frage, ob das Gebeverhalten von Gewaltstraftätern durch etablierte sozio-kontextuelle Determinanten, beispielsweise die erwartete Anzahl an Interaktionen und die Gruppenzugehörigkeit, adressierbar ist. Hierfür wurde die Gruppenzugehörigkeit mittels eines minimalen Gruppenparadigmas (Tajfel et al., 1971) eingeführt, bei dem die Teilnehmer lediglich erfuhren, dass Mitglieder der eigenen Gruppe sich sehr ähnlich und Mitglieder der fremden Gruppe sehr unähnlich waren. Des Weiteren wurde die Anzahl der erwarteten Interaktionen und damit die Möglichkeit zu Reziprozität experimentell manipuliert (eine Interaktion mit einem Mitspieler im ersten Teil des Experiments, mehrere Interaktionen im zweiten Teil des Experiments). Die Anzahl an geteilten Punkten wurde als kontinuierliches Maß für Kooperation verwendet (Van Lange & Kuhlman, 1994; Van Lange et al., 2002). Zudem wurden psychopathische Eigenschaften als potentieller Moderator dieser Zusammenhänge betrachtet. Die Ergebnisse dieser Studie lassen sich wie folgt zusammenfassen.

Es wurde betrachtet, inwiefern sich die Gruppenzugehörigkeit, die Anzahl an erwarteten Interaktionen sowie die experimentale Gruppe (inhaftierte Gewaltstraftäter und gesunde Kontrollen) sich auf das kooperative Verhalten auswirkte. Dabei zeigte sich ein Effekt der experimentalen Gruppe, der Gruppenzugehörigkeit und der Anzahl an erwarteten Interaktionen. Gegenüber der eigenen Gruppe verhielten sich die Teilnehmer im Schnitt wesentlich kooperativer. Zudem wurde im ersten Durchgang der wiederholten

Interaktion (also bei der Erwartung weiterer Interaktionen) in einem größeren Maß kooperiert im Vergleich zu einer einmaligen oder weiteren Interaktion in der wiederholten Bedingung. Des Weiteren zeigte sich, dass sich Gewaltstraftäter im Vergleich zu gesunden Kontrollen weniger Geldeinheiten mit ihren Mitspielern teilten, das heißt diese Individuen verhielten sich über alle Bedingungen hinweg weniger kooperativ. Zusätzlich wurde getestet, wie sich das kooperative Verhalten im Laufe des Experiments entwickelte. Gerade in der wiederholten Bedingung hätte es sein können, dass die Teilnehmer den computergenerierten Algorithmus durchschauen und dadurch ihr Verhalten verfälscht wird. Es konnte jedoch gezeigt werden, dass es lediglich in der ersten Interaktion (der wiederholten Bedingung) zu einem leichten Anstieg der Kooperation kam, während das kooperative Verhalten in der zweiten Interaktion stabil blieb. Diese Ergebnisse waren bei beiden Experimentalgruppen identisch. Anschließend wurde der Einfluss von psychopathischen Eigenschaften (gemessen mittels einer deutschen Version der Hare Self-Report Psychopathy Scale III) auf die Kooperationsbereitschaft untersucht. Hier zeigte sich bei der Gewaltstraftätergruppe, dass vor allem die Subskala kriminelle Tendenzen sowie im Trend die Subskala erratischer Lebensstil sowie der Gesamtscore negativ mit der Kooperationsbereitschaft in einfachen Interaktionen assoziiert waren. Zudem fanden wir in der Kontrollgruppe einen Zusammenhang zwischen höheren Werten auf der Subskala gefühlloser Affekt und niedrigeren Kooperationsraten mit Mitgliedern in der fremden Gruppe. Es ist anzumerken, dass diese Zusammenhänge nach einer Bonferroni Korrektur für multiples Testen nicht mehr signifikant waren.

Zusammenfassend zeigen die Ergebnisse dieser Studie, dass inhaftierte Gewaltstraftäter trotz einer generell geringeren Kooperationsbereitschaft durchaus in der Lage waren ihr Gebeverhalten an die Aufgabencharakteristika auf dieselbe Art und Weise wie gesunde Kontrollen anzupassen. Die Kooperationsbereitschaft der Gewaltstraftäter wurde durch die Anzahl an erwarteten Interaktionen sowie der Gruppenzugehörigkeit des Mitspielers in gleichem Maße wie bei gesunden Kontrollen beeinflusst. Dieser Befund spricht für die Existenz von äquivalenten psychologischen Mediatoren in beiden Gruppen. Zusätzlich konnte diese Studie belegen, dass sich bei antisozialen Individuen ein Minimalgruppen Paradigma umsetzen lässt – dies wurde

bisher in noch keiner wissenschaftlichen Arbeit belegt. Zudem legen unsere Daten die Vermutung nahe, dass psychopathische Eigenschaften bei Gewaltstraftätern mit sowohl strategischerem wie auch profit-maximierendem Verhalten einhergehen.

2.3 Studie 3: „Effects of group affiliation on neural signatures of fairness norm violations in antisocial violent offenders“

In Studie 2 konnte gezeigt werden, dass Menschen mit antisozialen Merkmalen weniger kooperativ trotz eines grundlegenden Regelverständnisses sind. Diese Studie vertauscht nun die vorigen Rollen und betrachtet antisoziale Gewaltstraftäter im Vergleich zu gesunden Kontrollen in der Rolle des Rezipienten im Ultimatum Game. Es war dabei von besonderem Interesse wie Individuen, die durch hochfrequente Regelverletzungen charakterisiert werden können, auf Fairnessnormverletzungen von anderen gegenüber ihnen selbst reagieren. Wir wollten herausfinden, ob die Entscheidungsfindung in einem solchen Kollektiv rein rational oder durch soziale Normen motiviert sind und ob sozio-kontextuelle Faktoren wie die Gruppenzugehörigkeit des Mitspielers bei der Entscheidungsfindung eine Rolle spielen. Die Teilnehmer erhielten im Ultimatum Game entweder faire Angebote (5:5) oder unfaire Angebote (1:9 und 3:7) von Mitspielern aus der eigenen oder fremden Gruppe. Darüber hinaus wurde die verhaltensbezogene Messung durch das Erfassen von elektrophysiologischen Reaktionen auf diese Fairnessnormverletzungen ergänzt. Die Ergebnisse der Studie lassen sich wie folgt zusammenfassen.

Antisoziale Gewaltstraftäter und gesunde Kontrollen nahmen fair und unfaire Angebote gleich häufig an, dabei wurden in beiden Gruppen faire Angebote häufiger akzeptiert. Dieser Befund gilt auch für die Manipulation der Gruppenzugehörigkeit: Beide Gruppen nahmen im gleichen Maß Angebote von Mitgliedern der eigenen und fremden Gruppe an, mit einer starken Bevorzugung der eigenen Gruppe. Zudem konnten die typischen elektrophysiologischen Reaktionen von gesunden Individuen auf Fairnessverletzungen in einem Gruppenkontext repliziert werden. Kontrollen zeigten eine wesentlich stärkere *Medial Frontal Negativity* (MFN) auf unfaire im Vergleich zu fairen Angeboten der eigenen Gruppe, während dieser Kontrast weniger negativ bei Angeboten der fremden Gruppe war. Im Gegensatz zu den behavioralen Ergebnissen, zeigten die antisozialen Gewaltstraftäter eine attenuierte elektrophysiologische Reaktion. Es zeigte sich kein

Unterschied zwischen Angeboten aus der eigenen und fremden Gruppe. Möglicherweise könnte die Reaktion auch genau andersherum im Vergleich zu den Kontrollen sein, das heißt antisoziale Personen zeigen bei Angeboten der Fremdgruppe eine stärkere MFN. Dieser Befund war jedoch statistisch nicht signifikant. Zusätzlich wurden in dieser Untersuchung die Ergebnisse nicht in einem signifikanten Maß von psychopathischen Eigenschaften beeinflusst. Dennoch konnten wir eine Tendenz für einen Zusammenhang zwischen einer stärkeren Attenuierung der MFN Reaktion und niedrigeren Akzeptanzraten für faire Angebote mit höheren psychopathischen Eigenschaften bei antisozialen Gewaltstraftätern beobachten.

Insgesamt liefert diese Studie Hinweise für Beeinträchtigungen in der Verarbeitung von sozialen Regelverletzungen auf einer elektrophysiologischen Ebene bei Menschen mit antisozialer Persönlichkeitsstörung, während wir für diese Individuen ähnliche Verhaltensmuster im Vergleich zu gesunden Kontrollen finden.

3. Diskussion

3.1 Determinanten sozialer Entscheidungsfindung

Das Ziel der hier vorgelegten Dissertation war die Untersuchung verschiedenster Facetten sozialer Entscheidungsfindung. In der folgenden Diskussion werden zunächst die Befunde zu gesunden Individuen diskutiert und in die bisherige Literatur eingeordnet. Wie in der Einleitung beschrieben, galt in der Ökonomie lange das Primat des Homo oeconomicus. Der Mensch wurde hier als ein Individuum beschrieben, das ausschließlich den eigenen Gewinn vergrößern möchte und dabei ohne wesentliches Interesse gegenüber anderen handelt (Persky, 1995). Es wurde bereits erwähnt, dass diese Sichtweise zunehmend widerlegt wurde und dem Menschen durchaus prosoziales und weniger rationales Verhalten zugetraut wurde (Balliet et al., 2014; Eisenberg & Miller, 1987; Gintis, 2000).

Die Daten dieser Dissertation bestätigten letztere Aussage. In einem Standard Dictator Game (Studie 1) wurde mit 30% der zur Verfügung stehenden Ressourcen ein wesentlicher Teil abgegeben. Ein rationaler Agent würde in einer solchen Interaktion, ohne Konsequenzen oder Vergeltung des anderen Spielers, keine oder sehr wenige Ressourcen teilen. Dennoch entspricht der referierte Befund den Ergebnissen aus einer umfangreichen Metaanalyse zum Dictator Game (Engel, 2011). Damit leisten die vorgestellten Daten einen weiteren Beitrag zur Replikation des Befundes, dass sich gesunde Individuen im Durchschnitt altruistisch und prosozial anderen gegenüber verhalten und eben keine rationalen Spieler sind. Die Ergebnisse zu den gesunden Teilnehmern aus Studie 2 widersprechen ebenfalls den Annahmen des Homo oeconomicus: In einem gewöhnlichen Give Some Dilemma gibt es zwar nicht eine klar rationale Strategie, dennoch gab es in dieser Studie eine Bedingung, bei der die Teilnehmer einmalig mit einem Mitspieler interagierten und dadurch das eigene Verhalten keinen Einfluss auf weitere Durchgänge mit anderen Mitspielern hatte. Selbst in dieser Bedingung waren gesunde Individuen bereit über vier von zehn möglichen Geldeinheiten abzugeben. Dieser Befund repliziert ebenfalls vorige Studien, die zu gleichen Ergebnissen kommen und damit den Homo oeconomicus in Frage stellen (Van Lange et al., 2011; Van Lange et al., 2002). Abschließend können die Ergebnisse aus Studie 3 auf dieselbe Art und Weise interpretiert werden. Hier spielten die Teilnehmer

als Responder in einem Ultimatum Game, wobei ein rationaler Spieler jedes noch so kleine Angebot annehmen würde. Zwar lehnten gesunde Teilnehmer in dieser Studie unfaire Angebote häufiger als faire Angebote ab, dennoch wurden unfaire Angebote von Mitgliedern der eigenen Gruppe in der Hälfte der Fälle und Angebote von Fremdgruppenmitgliedern in einem Drittel der Fälle akzeptiert. Somit sprechen auch diese Daten gegen ein rationales Spielen des Ultimatum Games. Die Replikation vorheriger Studien mit gleichem Ergebnis stärkt diese Aussage zusätzlich (Camerer, 2003; Sanfey et al., 2003; Yamagishi et al., 2009). Zusammenfassend konnten im Rahmen dieser Dissertation über verschiedene Paradigmen hinweg und mit unterschiedlichen (gesunden) Stichproben einerseits typische Befunde aus dem Forschungsfeld der sozialen Entscheidungsprozesse repliziert werden und damit andererseits die Annahme des Homo oeconomicus weiter in Frage gestellt werden.

3.1.1 Empathie

Als eine zentrale Determinante sozialer Entscheidungsfindung wurde in dieser Dissertation die situationelle Empathie sowie die Persönlichkeitseigenschaft der Empathiefähigkeit untersucht. Die Daten aus Studie 1 bestätigen die Empathie-Altruismus Hypothese und erweitern deren Gültigkeit in ökonomischen Interaktionen. Bei dieser Hypothese wird angenommen, dass eine altruistische Motivation durch empathische Gefühle einer hilfsbedürftigen Person gegenüber ausgelöst werden (Batson, 1990). Zunächst konnte in dieser Studie bestätigt werden, dass Menschen in einem gewöhnlichen Dictator Game und in einer Spielsituation ohne Empathie-Induktion übereinstimmend mit Fairnessnormen einen substantiellen Anteil der zur Verfügung stehenden Ressourcen abgeben (zwischen 30% und 40%). Darauf aufbauend zeigte die Empathie-Induktion die vorhergesagte Wirkung: Die Teilnehmer waren im Schnitt bereit über 70% des Gesamtbetrages an hilfsbedürftige andere abzugeben. Dieses wesentlich altruistische Verhalten konnte zudem durch einen Anstieg von empathischen Gefühlen erklärt werden. Mit anderen Worten war es in diesem Experiment möglich, Fairnessnormen durch eine gezielte Empathie-Induktion zu übertreffen. Somit konnte hier erstmals die Annahme, dass Empathie eine zentrale Determinante altruistischen Verhaltens gegenüber Notleidenden ist (de Waal, 2008), auf ökonomische Interaktionen im Sinne des Dictator Games erweitert werden.

Es ist anzumerken, dass bisherige Studien eine geringere Bereitschaft zu altruistischem Verhalten gegenüber Wohltätigkeitsorganisationen (Eckel & Grossman, 1996) oder erfundenen Personen als Empfänger (Telle & Pfister, 2012) beobachteten. Dieser Unterschied und der damit verbundene stärkere Altruismus können dadurch erklärt werden, dass in der vorgestellten Studie die Entscheidungen der Teilnehmer direkt mit ihrem realen Gewinn zusammenhing und sie mit visuellen Darstellungen der leidenden Empfänger konfrontiert waren. Zudem konnte Studie 1 zeigen, dass der Grad der Empathie gegenüber einer bestimmten Person situationsabhängig variiert und altruistisches Verhalten im wirtschaftlichen Kontext motiviert. Damit werden frühere Erkenntnisse zum Verhältnis von Empathie als Persönlichkeitseigenschaft und Altruismus gegenüber anonymen Fremden (Edele, Dziobek, & Keller, 2013) sowie die Erforschung der Rolle von Mitgefühl als Motivator für prosoziales Verhalten bei weniger religiösen Teilnehmern (Saslow et al., 2013) erweitert. Darüber hinaus liefern diese Ergebnisse eine Erklärung für die Beobachtung, dass wenn man Opfern ein Gesicht gibt, dies zu altruistischem Verhalten führt (Small & Loewenstein, 2003). In Studie 1 konnte gezeigt werden, dass das Ausmaß an Empathie gegenüber einer anderen Person stark voraussagt, inwiefern Menschen bereit sind in einer ökonomischen Situation zu teilen und damit altruistisches Verhalten zu zeigen. Das Sehen des Gesichtes eines Opfers könnte Empathie und damit Altruismus hervorrufen.

Des Weiteren sagten in Studie 1 die selbsteingeschätzten empathischen Gefühle einen großen Teil des altruistischen Verhaltens vorher, während empathische Persönlichkeitseigenschaften weniger aussagekräftig für dieses Verhalten waren. Für diese Ergebnisse lassen sich in der Literatur widersprüchliche Studien finden, so wurde unter anderem ein signifikanter Zusammenhang zwischen Empathie als Persönlichkeitseigenschaft und dem Verhalten in einem gewöhnlichen Dictator Game gefunden (Edele et al., 2013). Dennoch gibt es auch metaanalytische Evidenz, die einen stärkeren Zusammenhang von prosozialem Verhalten und situativer Empathie im Vergleich zu empathischen Eigenschaften betont (Eisenberg & Miller, 1987). Die geringeren Zusammenhänge für empathische Persönlichkeitseigenschaften können möglicherweise durch die Heterogenität der Fragebogenmaße erklärt werden (Eisenberg & Miller, 1987). Zusammenfassend kann aus Studie 1 geschlossen werden,

dass man um altruistisches Verhalten zu fördern – unabhängig davon ob für Wohltätigkeitsorganisationen, Flüchtlinge oder in anderen ökonomischen und politischen Kontexten – auf jeden Fall die empathischen Gefühle einer Person für den speziellen Empfänger ansprechen sollte.

3.1.2 Reziprozität

Als weitere wichtige Determinante auf die soziale Entscheidungsfindung wurde im Rahmen dieser Dissertation die sozio-kontextuellen Faktoren der Zugehörigkeit des Interaktionspartners sowie die Anzahl an erwarteten Interaktionen mit diesem untersucht. Bezüglich letzterer Variable konnte in Studie 2 gefunden werden, dass die Antizipation von wiederholten Interaktionen im Vergleich zu einfachen Interaktionen mit dem gleichen Mitspieler zu einer Steigerung prosozialem Verhaltens in Form von Kooperation führt. Dieses Ergebnis repliziert einen Vorbefund mit demselben Paradigma (Van Lange et al., 2011). In der Literatur wird davon ausgegangen, dass die Bereitschaft zu mehr Kooperation in wiederholten Interaktionen sich durch Vorteile dieses prosozialen Verhaltens sowie einer potentiellen Bestrafung bei unkooperativen Verhaltensweisen und der daraus entstehenden Reziprozität erklärt (Axelrod & Hamilton, 1981).

Betrachtet man das Give Some Dilemma näher, so wird ersichtlich wie diese Reziprozität entsteht und verhaltenswirksam wird. Dieses Paradigma zeichnet sich im Fall von einfachen und wiederholten Interaktionen durch Aspekte des Dictator Games (Fehlen von Konsequenzen in der einfachen Interaktion) und dem Ultimatum Games (Möglichkeit vom Mitspieler bestraft zu werden in der wiederholten Interaktion) aus. Interessanterweise finden sich analoge Ergebnisse für diese beiden Paradigmen mit höheren Angeboten im Ultimatum Game (Camerer, 2003; Sanfey et al., 2003; Yamagishi et al., 2009) im Vergleich zum Dictator Game (Engel, 2011). Während im Ultimatum Game ein Mitspieler ausschließlich bestrafen kann und das auch tut, wenn er unfaire Angebote bekommt (Bolton & Zwick, 1995), gibt es im Give Some Dilemma auch die Möglichkeit einen Mitspieler für kooperatives Verhalten mit einem eigenen kooperativen Angebot zu belohnen. Dadurch können sowohl Bestrafung von unkooperativen Verhaltensweisen wie auch Belohnung von Kooperation zu den Unterschieden zwischen einfachen und wiederholten Interaktionen in Studie 2 geführt

haben. Weitere Studien könnten mit einem modifizierten Paradigma durch eine zusätzliche Option zur Bestrafung von Angeboten überprüfen, welcher der beiden Mechanismen zu den referierten Ergebnissen führt.

Für den Befund der höheren Kooperationsbereitschaft in den wiederholten Interaktionen könnte zudem die Alternativerklärung gelten, dass die Teilnehmer während des Spiels lernten und das vorab definierte Verhalten des Mitspielers „durchschauten“. Der computerprogrammierte Mitspieler machte stets zuerst ein faires Angebot und antwortete danach in der zweiten Interaktion nach dem so genannten „tit-for-tat“ Prinzip („Wie du mir, so ich dir“) (Axelrod & Hamilton, 1981). Daher besteht die Möglichkeit, dass die Teilnehmer diesen Algorithmus im Laufe des Experiments verstanden und sich im ersten Angebot sehr kooperativ verhielten, um dadurch ein besseres zweites Angebot und somit einen höheren Gewinn zu bekommen (Burton-Chellew, Nax, & West, 2015). Tatsächlich war ein Anstieg der Kooperationsraten in der ersten wiederholten Interaktion im Laufe des Experiments zu beobachten. Gegen ein Durchschauen der experimentellen Programmierung und einer damit verbundenen Verbesserung des persönlichen Gewinns sprechen jedoch die stabilen Kooperationsraten in der zweiten wiederholten Interaktion. Zudem wurde auch das Verhalten durch die Gruppenzugehörigkeit in der wiederholten Interaktion beeinflusst – wären die Teilnehmer also ausschließlich auf eine Gewinnmaximierung aus gewesen, dann sollte dieser Faktor keine Rolle spielen. Der leichte Anstieg bezüglich kooperativen Verhaltens in der ersten Interaktion kann daher besser durch den Aufbau von Vertrauen in die Reziprozität der Interaktion erklärt werden.

3.1.3 Gruppenzugehörigkeit

Als weiterer in dieser Dissertation betrachteter sozio-kontextueller Einflussfaktor auf die soziale Entscheidungsfindung ist die Gruppenzugehörigkeit des Interaktionspartners zu nennen. In Studie 2 und Studie 3 konnten der typische in der Literatur beschriebene Effekt des Eigengruppen-Bias bestätigt werden (Brewer, 1979). Studie 2 konnte diesen Befund erstmals in einem Give Some Dilemma belegen, indem gezeigt werden konnte, dass sich Menschen gegenüber Mitgliedern der eigenen Gruppe kooperativer im Vergleich zu Fremdgruppenmitgliedern verhielten. In Studie 3 spielten die Teilnehmer als Responder ein Ultimatum Game und auch hier zeigte sich eine deutliche

Bevorzugung der eigenen Gruppe: In diesem Experiment nahmen die Teilnehmer Angebote der eigenen Gruppe wesentlich häufiger an im Vergleich zu Angeboten von Fremdgruppenmitgliedern, wobei dieses Ergebnis unabhängig von der Fairness der Angebote war. Es konnte somit über verschiedene Experimente mit unterschiedlichen Methoden der Eigengruppen-Bias repliziert und auf ein neues Paradigma erweitert werden.

Zudem konnte ebenfalls bestätigt werden, dass diese Gruppenprozesse auf einer sehr basalen Ebene ablaufen: In beiden Experimenten wurde ein Minimal-group Paradigma angewendet, bei dem die Teilnehmer (ohne ihr Wissen) zufällig einer Gruppe zugeteilt wurden, nicht miteinander kommunizieren konnten und keine anderen Mitglieder der eigenen oder fremden Gruppe kannten (Brewer, 1979, 1999; Tajfel et al., 1971; Tajfel & Turner, 1979). Es ist hervorzuheben, dass diese Art der experimentellen Manipulation bisher noch nicht in einem Give Some Dilemma angewandt wurde und daher sind die Befunde aus Studie 2 als Erweiterung der Literatur auf ein neues Paradigma sowie die Ergebnisse aus Studie 3 als Replikation voriger Studien zu verstehen. Der Befund, dass es zu einem noch stärkeren Eigengruppen-Bias kommt, wenn das eigene Ergebnis im Sinne der Reziprozität vom Verhalten des Mitspielers abhängt (Balliet et al., 2014), konnte in Studie 2 nicht bestätigt werden, da es hier keinen Zusammenhang zwischen den beiden Faktoren gab. Die Daten dieser Dissertation sprechen somit für eine Unabhängigkeit dieser wichtigen sozio-kontextuellen Faktoren.

Es wird in der Literatur davon ausgegangen, dass der beschriebene Eigengruppen-Bias vor allem aus dem Wunsch entsteht, positive Verhältnisse in der eigenen Gruppe aufrechtzuerhalten und zu stärken (Brewer, 1999). Es ist jedoch auch anzumerken, dass in beiden referierten Studien gesunde Individuen mit inhaftierten Gewaltstraftätern verglichen wurden. Dieser Umstand könnte zumindest teilweise zu den starken Gruppeneffekten beigetragen haben. In beiden Studien erhielten die Teilnehmer die Informationen, dass sie in eine von zwei Gruppen aufgrund der Beantwortung eines vorigen Fragebogens zugeteilt werden. Dieser Fragebogen hatte das Ziel Psychopathie und Aggressivität zu messen. Vor diesem Hintergrund könnte es also sein, dass vereinzelte Teilnehmer die (möglicherweise aggressiven, psychopathischen) inhaftierten Individuen als Fremdgruppe und die nicht aggressiven, nicht psychopathisch, nicht

verurteilten Teilnehmer als Eigengruppe wahrgenommen haben. Eine Limitation beider Studien besteht in dieser eher unspezifischen Zuteilung der Gruppen. Es wurde zudem nicht überprüft welche Eigenschaften der Mitspieler zu einer Identifikation der Gruppenmitglieder führte. Zukünftige Studien sollten daher untersuchen, welche Charakteristika der eigenen und der fremden Gruppe zugeschrieben werden. Dennoch gilt es zu betonen, dass gerade die Stabilität der Ergebnisse zur Gruppenmanipulation über zwei Experimente sowie identische Ergebnisse aus einer langen Forschungstradition gegen eine Konfundierung der Ergebnisse durch die beiden experimentellen Gruppen (inhaftierte Gewaltstraftäter und gesunde Kontrollen) sprechen.

3.1.4 Affektive Korrelate

In Studie 3 wurden die Untersuchungen von Gruppenprozessen im Ultimatum Game um die Messung elektrophysiologischer Daten erweitert. Die Fragestellung hier war, wie sich die Evaluation fairer und unfairer Angebote durch Mitglieder der eigenen oder fremden Gruppe auf einer neuralen Ebene beziehungsweise bei einer Elektroenzephalografie (EEG) abbilden. Das Forschungsdesign wurde hierfür an eine vorige Studie angelehnt (Wang et al., 2017), deren Ergebnisse repliziert und um Aussagen zur Gruppe der inhaftierten Gewaltstraftäter erweitert wurden. Wie auch bei Wang et al. (2017) zeigte sich für gesunde Individuen, dass die Gruppenzugehörigkeit frühe Prozesse der Fairnessevaluation im Sinn einer veränderten *Medial Frontal Negativity* (MFN) beeinflusst. Die MFN ist ereigniskorreliertes Potential und eine negative Deflektion, die unfairen im Vergleich zu fairen Angeboten nachfolgt und daher eine Art affektive Sensibilität gegenüber Verletzungen von Fairnesserwartung darstellt (Boksem & De Cremer, 2010; Boksem et al., 2006; Gehring & Willoughby, 2002; Hewig et al., 2011; Masaki et al., 2006; Pfabigan et al., 2011; Potts et al., 2006; Sato et al., 2005). Wie erwartet war die Ausprägung der MFN größer wenn Eigengruppenmitglieder (im Vergleich zu Fremdgruppenmitgliedern) die Angebote machten. Zu diesem Ergebnis kam auch die Studie von Wang et al. (2017). Die Autoren schlagen als Erklärung dieser Ergebnisse vor, dass die MFN durch gruppenbasierte Fairnesserwartungen beeinflusst wird. Gesunde Individuen scheinen faire Angebote von Eigengruppenmitgliedern zu erwarten, daher könnte ein unfaires Angebot eine stärkere Negativierung der MFN

hervorrufen. Es ist anzumerken, dass die gesunden Teilnehmer in Studie 3 und bei Wang et al. (2017) dennoch häufiger (faire und unfaire) Angebote von Eigengruppenmitgliedern annehmen. Daraus kann geschlossen werden, dass sich die Teilnehmer trotz dieser Fairnesserwartung dazu entscheiden sich im Sinne der sozialen Norm der Eigengruppenbevorzugung zu verhalten. Wang et al. (2017) vermuten, dass die Versuchsteilnehmer sich so verhalten, um die Gruppenintegrität zu erhalten und stabilisieren, wenn einzelne Gruppenmitglieder gegen Normen verstoßen.

3.2 Soziale Entscheidungsfindung bei antisozialen Individuen

Zu den referierten Befunden zu gesunden Individuen wurde zudem die Persönlichkeitseigenschaft Dissozialität im Sinne der antisozialen Persönlichkeitsstörung betrachtet. Im Rahmen dieser Dissertation kann die Frage aufgeworfen werden, ob Personen aus dem antisozialen Formkreis wie postuliert erhebliche Defizite in der sozialen Entscheidungsfindung aufweisen. Im klinisch-psychotherapeutischen Sinn wird für die Vergabe der Diagnose der antisozialen Persönlichkeitsstörung verlangt, dass ein Individuum tiefgreifende Muster von Missachtung und Verletzungen der Rechte anderer aufweist (American Psychiatric Association, 2013). Zudem gibt es Kriterien wie das Versagen sich an Normen anzupassen, andere zu betrügen, Verantwortungslosigkeit sowie fehlende Reue gegenüber den Gesetzesbrüchen. Es wäre ungewöhnlich überhaupt Anzeichen von prosozialem Verhalten oder intakten Entscheidungsprozessen anzunehmen, falls diese Kriterien bei einem Individuum zutreffen. Zudem wurden in der bisherigen Forschungsliteratur Zusammenhänge zwischen antisozialem Verhalten beziehungsweise Kriminalität und Defizite bezüglich empathischer Fähigkeiten vermutet. Genauer wurde angenommen, dass die Wahrnehmung von Distress bei anderen durch den Aggressor zu einer Hemmung von gewalttätigem Verhalten führt (Miller & Eisenberg, 1988). Empirisch zeigt sich ein Zusammenhang zwischen Straffälligkeit, vor allem Gewaltstraftaten, und Defiziten bezüglich der Empathie (Jolliffe & Farrington, 2004). Da die Ergebnisse von Studie 1 eindeutige Hinweise auf einen Zusammenhang zwischen empathischen States und Traits mit prosozialem Verhalten (in diesem Fall Altruismus) liefern, kann weiter angenommen werden, dass gerade antisoziale Gewaltstraftäter Schwierigkeiten bezüglich prosozialem Verhalten haben müssten.

3.2.1 Abweichendes Verhalten durch strategisches und rationales Handeln

Betrachtet man die bisherige Literatur zu dieser Fragestellung, so wird deutlich, dass das Forschungsfeld bisher zu noch keiner klaren Antwort gekommen ist. Über verschiedene Studien hinweg ist lediglich der Befund bezüglich des Gebeverhaltens im Dictator Game konsistent. Hier konnte gezeigt werden, dass bei Studierenden (primär) psychopathische Eigenschaften mit geringeren Altruismus in diesem Spiel korreliert waren (Berg, Lilienfeld, & Waldman, 2013). Eine Studie mit inhaftierten Straftätern kam ebenfalls zum Schluss, dass primäre Psychopathen im Vergleich zu Nicht-Psychopathen und sekundären Psychopathen weniger Altruismus im Dictator Game zeigen (Koenigs, Kruepke, & Newman, 2010). Es ist anzumerken, dass die beiden zuletzt genannten Gruppen mit über 40% abgegebener Geldeinheiten sich durchaus in einem unauffälligen Rahmen bewegten – trotz antisozialer Eigenschaften. Weitere Studien kamen zu ähnlichen Resultaten: Es konnte gezeigt werden, dass Gefangene im Vergleich zu gesunden Personen im gleichen Alter weniger im Dictator Game gaben (Chmura, Engel, & Englerth, 2017; Mayer, Jusyte, Klimecki-Lenz, & Schönenberg, 2018; Mayer, Rauss, Pourtois, Jusyte, & Schönenberg, 2018). Droht also keine unmittelbare Strafe, dann verhalten sich antisoziale Individuen weniger prosozial.

Diese Befunde lassen sich durch Untersuchungen mit anderen Paradigmen ergänzen, welche weitere Facetten der Entscheidungsfindung bei antisozialen Individuen untersuchen. Bezogen auf prosoziales Verhalten in Form von Kooperation kommen verschiedene Studien mit studentischen Stichproben zum Schluss, dass psychopathische Merkmale mit geringerer Kooperationsbereitschaft im Prisoners Dilemma assoziiert sein könnten (Berg et al., 2013; Curry, Chesters, & Viding, 2011; Gervais, Kline, Ludmer, George, & Manson, 2013; Johnston, Hawes, & Straiton, 2014; Rilling et al., 2007). Eine vielbeachtete Studie konnte zudem bei inhaftierten Psychopathen weniger kooperatives Verhalten im Vergleich zu Kontrollen feststellen (Mokros et al., 2008). Dennoch kam eine weitere Studie mit inhaftierten Straftätern zu einem gegenteiligen Ergebnis, nämlich dass Straftäter im Vergleich zu Kontrollen in gleichem Maß kooperatives Verhalten zeigten (Khadjavi & Lange, 2013). Es ist anzumerken, dass diese Studie weibliche Gefangene testete. Vor dem Hintergrund, dass es Geschlechterunterschiede bezüglich kooperativem Verhalten gibt, Frauen

kooperieren initial mehr im Vergleich zu Männern, wobei diese Unterschiede im Verlauf einer Interaktion verschwinden (Ortmann & Tichy, 1999), ist eine Vergleichbarkeit dieser Studien nicht unbedingt gegeben. Es wurden ebenfalls unterschiedliche experimentelle Designs verwendet (einfache vs. wiederholte Interaktionen). Auch in diesem Paradigma gibt es keine unmittelbare Strafe für unkooperatives Verhalten. Zwar war beispielsweise das Experiment von Mokros und Kollegen (2008) so programmiert, dass der Computeralgorithmus nach mehrfachem unkooperativen Verhalten ebenfalls unkooperativ reagierte (so genannte „tit-for-two-tats“ Strategie, eine Abwandlung der „tit-for-tat“ Strategie, zu Deutsch „wie du mir, so ich dir“), dennoch ermöglichte dieses Verfahren hochfrequentes unkooperatives Verhalten ohne negative Konsequenzen oder Vergeltung von anderen.

Insgesamt kann aus diesen Studien geschlossen werden, dass Individuen aus dem antisozialen Formkreis in Situationen, bei denen keine Strafe droht, sich weniger prosozial verhalten und eher den eigenen Vorteil suchen. Ein Teil der Befunde aus Studie 2 bestärkt diese Vermutung, da auch hier inhaftierte Gewaltstraftäter sich im Schnitt weniger kooperativ verhielten. Es ist anzumerken, dass in dem hier verwendeten Paradigma Konsequenzen ebenfalls eine wesentliche Rolle spielten. Interagierten die Teilnehmer nur einmalig mit einer Person, wurde weniger kooperiert im Vergleich zu einer wiederholten Bedingung. Dieses Verhalten war jedoch nicht spezifisch für die inhaftierte Gruppe.

Anknüpfend an den Aspekt der Konsequenzen von Verhalten, lohnt sich der Blick auf Studien mit dem Ultimatum Game als Paradigma. Hier zeigen sich in der Rolle des Gebenden keine wesentlichen Defizite bei antisozialen Individuen. In zwei Studien verhielten sich antisoziale und psychopathische Gefangene ähnlich wie gesunde Kontrollen (Koenigs et al., 2010; Mayer, Rauss, et al., 2018). Eine weitere Studie mit einer studentischen Stichprobe fand zudem verschiedene Zusammenhänge zwischen psychopathischen Eigenschaften und dem Gebeverhalten im Ultimatum Game. Unter anderem fanden die Autoren, dass höhere Angebote mit höherem machiavellistische Egozentrismus assoziiert waren (Curry et al., 2011). In der Literatur wird aus diesen Ergebnissen oft geschlossen, dass Individuen aus dem antisozialen und psychopathischen Spektrum möglicherweise ein stark strategisch motiviertes Verhalten

zeigen (Berg et al., 2013; Curry et al., 2011; Mayer, Rauss, et al., 2018). Das bedeutet, es wird unterschieden ob ein Verhalten bestraft wird (wie im Ultimatum Game) oder nicht (wie beispielsweise im Dictator Game) und so das Verhalten auf einen möglichst großen (eigenen) Gewinn ausgerichtet ist. Es ist anzumerken, dass trotz dieser referierten Untersuchungen die Befunde bei weitem nicht so stark sind, wie man vor dem Hintergrund der vorigen Überlegungen annehmen müsste. Zudem zeigen auch gesunde Personen diese Verhaltenstendenz: Während im Ultimatum Game im Durchschnitt 50% der Ressourcen abgegeben werden (Camerer, 2003; Sanfey et al., 2003; Yamagishi et al., 2009), sind es im Dictator Game lediglich 30% im Durchschnitt (Engel, 2011). Die Richtung des Verhaltens scheint also bei antisozialen Individuen und Gesunden dieselbe zu sein, wobei die Entscheidungsprozesse der klinischen Population möglicherweise einer größeren Varianz unterliegen.

Die Ergebnisse von Studie 2 lassen zusätzlich an einer klaren Schlussfolgerung zweifeln: Wir fanden zwar, dass inhaftierte Gewaltstraftäter über alle Bedingungen hinweg weniger prosoziales Verhalten zeigten und dennoch nutzte diese Gruppe die unterschiedlichen Bedingungen nicht übermäßig aus. In anderen Worten, zeigte die gesunde Kontrollgruppe das exakt selbe Verhalten (weniger Punkte in einer einfachen Interaktion abgeben im Vergleich zu Interaktionen, in denen weitere Begegnungen erwartet werden) mit dem Unterschied insgesamt etwas mehr zu geben. Dieses Ergebnis widerspricht den Befunden aus einer Stichprobe von Studenten, bei der psychopathische Eigenschaften mit einer stärkeren und damit strategischeren Berücksichtigung von Hinweise für eine erneute Interaktion in der Zukunft verbunden waren (Gervais et al., 2013). Die Autoren dieser Studie interpretierten ihre Ergebnisse ebenfalls als Hinweis für strategische Überlegungen bei Psychopathen. Schlussendlich fanden auch wir in unserer Studie einen Zusammenhang zwischen psychopathischen Eigenschaften und geringeren Abgaben in einfachen Interaktionen. Aufgrund der korrelativen Natur dieser Aussagen bleiben endgültige Schlüsse dennoch lediglich spekulativ. Es könnte möglich sein, dass gerade diese psychopathischen Eigenschaften zu einem rationalen und strategischen Verhalten führen. Dennoch ist auch hier die Befundlage als unbefriedigend einzuschätzen, da die Stichproben sich sehr unterscheiden (Studenten vs. Inhaftierte) und auch die Befunde zu psychopathischen

Eigenschaften durchaus heterogen sind: In der Studie von Gervais und Kollegen (2013) sind primäre psychopathische Eigenschaften der beste Prädiktor, während unsere Daten ausschließlich Zusammenhänge für sekundäre Psychopathie finden.

3.2.2 Einfluss sozio-kontextueller Determinanten

Wie gestört sind nun Entscheidungsprozesse bei antisozialen Individuen? Die bisher referierte Literatur sowie die vorgestellten eigenen Ergebnisse lassen noch keinen endgültigen Schluss zu. Wäre das Verhalten tatsächlich strategischer, so sollte es keine Rolle spielen, mit wem interagiert wird, so lange der eigene Vorteil gewahrt wird. Interessanterweise finden sich dennoch Studien mit Studenten, die Hinweise auf eine erhöhte Eigengruppenbevorzugung bei erhöhten psychopathischen Eigenschaften finden (Arbuckle & Cunningham, 2012; Gillespie, Mitchell, Johnson, Dawson, & Beech, 2013). Studie 2 findet zwar keine stärkere Eigengruppenbevorzugung, dafür ist diese Präferenz jedoch im selben Maß wie bei der gesunden Kontrollgruppe vorhanden. Die Unterschiede in der Richtung dieser Verhaltenstendenz können durch methodische Aspekte erklärt werden (Stichprobe: Studenten vs. inhaftierte Gewaltstraftäter sowie Implementierung der Gruppenzugehörigkeit: Eigengruppe und unbekannte Personen vs. Eigengruppe vs. Fremdgruppe). Es bleibt jedoch festzuhalten, dass antisoziale Personen in einem Gruppenkontext in keiner Weise ein strategisches Verhalten zeigen und adäquat auf diesen sozio-kontextuellen Faktor reagieren.

Die Ergebnisse aus Studie 3 bestätigen dieses nicht-strategische Verhalten in Gruppenkontexten bei antisozialen Individuen: Wir fanden, dass inhaftierte Gewaltstraftäter faire und unfaire Angebote von Eigen- und Fremdgruppemitgliedern in gleichem Maß wie gesunde Kontrollen annahmen. Dabei wurden tendenziell mehr Angebote aus der eigenen Gruppe akzeptiert, was auch hier für eine Bevorzugung der eigenen Gruppe spricht. Somit konnte auch hier gezeigt werden, dass antisoziale Individuen keine defizitäre Repräsentation von Fairness aufweisen beziehungsweise nicht stärker gegen deren Verletzungen vorgehen. Dennoch ist die übrige Literatur bezüglich der Entscheidungsfindung in der Rolle des Empfängers im Ultimatum Game weniger eindeutig. Zwei Studien fanden keine Unterschiede zwischen inhaftierten

Personen und Kontrollen (Koenigs et al., 2010; Radke, Brazil, Scheper, Bulten, & De Bruijn, 2013), jedoch fand eine andere Studie grundsätzlich höhere Akzeptanzraten und damit ein strategischeres Verhalten bei inhaftierten Gewaltstraftätern (Mayer, Rauss, et al., 2018). Studien mit studentischen Stichproben kamen zu denselben widersprüchlichen Ergebnissen für Teilnehmer mit hohen psychopathischen Eigenschaften (Osumi & Ohira, 2010; Vieira et al., 2014), wobei in der Studie von Koenigs und Kollegen (2010) bei inhaftierten Straftätern psychopathische Eigenschaften mit niedrigeren Akzeptanzraten einher gingen. Schlussendlich kann auch aus dieser Studienlage kein strategisches Verhalten abgeleitet werden.

3.2.3 Neurophysiologische Aspekte

Ein zusätzlicher Parameter im Rahmen von sozialen Entscheidungsprozessen sind physiologische Messungen. Im Rahmen dieser Arbeit wird im Folgenden auf elektrophysiologische Korrelate der Elektroenzephalografie (EEG) eingegangen. Für solche Maße ist besonders das Ultimatum Game mit der Rolle als Empfänger von Angeboten geeignet, da so die Reaktionen und deren frühe Verarbeitungsprozesse auf beispielsweise Fairnessverletzungen beobachtet werden können. Ein vieluntersuchter Parameter in diesem Paradigma ist die so genannte *Medial Frontal Negativity* (MFN). Dieses ereigniskorrelierte Potential wird häufig als Index für eine affektive Evaluation negativer Outcomes (Boksem et al., 2006; Gehring & Willoughby, 2002; Masaki et al., 2006), Verletzung sozialer Normen (Boksem & De Cremer, 2010) oder Erwartungen verstanden (Pfabigan et al., 2011; Potts et al., 2006; Sato et al., 2005). Die MFN ist gerade deshalb für die Untersuchungen von Entscheidungsprozessen bei antisozialen Individuen von Bedeutung, weil sie im anterioren cingulären Cortex generiert wird. Dieser Teil des Gehirns ist bei Entscheidungsprozessen im Allgemeinen zentral involviert (Bush, Luu, & Posner, 2000; Eisenberger, Lieberman, & Williams, 2003; Sanfey et al., 2003; Shackman et al., 2011). Zudem weisen Individuen aus dem antisozialen Spektrum strukturelle und funktionale Veränderungen in dieser Region auf (Glenn, Raine, & Schug, 2009; Glenn, Yang, Raine, & Colletti, 2010; Kiehl, 2006; Kiehl et al., 2001; Koenigs, 2012; Ly et al., 2012; Rilling et al., 2007; Yang & Raine, 2009).

Erst zwei Studien untersuchten die MFN bei antisozialen Individuen. In einer ersten Untersuchung unserer Arbeitsgruppe wurde das Verhalten von inhaftierten

Gewaltstraftätern als Empfänger fairer und unfairer Angebote von Menschen oder vom Computer bei einem Ultimatum Game untersucht (Mayer, Rauss, et al., 2018). Dabei zeigten die inhaftierten Gewaltstraftäter eine abgeschwächte MFN, die darauf hindeutete, dass diese Individuen auf elektrophysiologischer Ebene weniger auf Fairnessverletzungen (unabhängig davon, ob das Angebot von einem Menschen oder vom Computer kam) reagierten. Die Daten aus Studie 3 bestätigen diesen Befund und erweitern ihn um den sozio-kontextuellen Faktor der Gruppenzugehörigkeit: In dieser Studie fanden wir ebenfalls eine attenuierte MFN bei inhaftierten Gewaltstraftätern als Reaktion auf unfaire und faire Angebote von Mitgliedern der eigenen oder einer fremden Gruppe. Genauer gesagt, unterschieden die Teilnehmer dieser Studie nicht zwischen fairen und unfairen Angeboten aus der eigenen oder fremden Gruppe. Im Gegensatz hierzu verarbeiteten die gesunden Kontrollen die Angebote der beiden Gruppen unterschiedlich. Zusammenfassend kann aus diesen beiden Studien geschlossen werden, dass frühe Verarbeitungsprozesse bei antisozialen Individuen in Entscheidungsprozessen verändert sein könnten.

Mayer et al. (2018) schlagen vor, dass diese Defizite ein rationaleres Verhalten bei antisozialen Individuen erklären könnten, da emotionale Reaktionen auf Erwartungs- und Normverletzungen weniger ausgeprägt zu sein scheinen. Da jedoch die Befunde zum Verhalten bei den betroffenen Individuen nicht auf ein rationaleres Verhalten hindeutet, scheinen andere Erklärungen notwendig zu sein. Zunächst könnten antisoziale Individuen ein grundsätzlich anderes (oder gestörtes) Verständnis von Normen haben. Diese Idee stimmt mit dem Befund überein, dass Studenten mit hohen psychopathischen Eigenschaften unfaire Angebote im Ultimatum Game als weniger unfair wahrnahmen (Vieira et al., 2014). Gegen diese Vermutung spricht jedoch, dass die Akzeptanzraten zwischen hoch und niedrig psychopathischen Studenten in dieser Studie gleich waren. Auch unsere Daten aus Studie 3 zeigen, dass antisoziale Individuen auf der Verhaltensebene ein gewöhnliches Normverständnis zeigen. Hätten diese Menschen andere Normen, dann müssten sie zumindest teilweise ein entsprechendes Verhalten in diesen Paradigmen zeigen. Es scheint wahrscheinlicher, dass antisoziale Individuen in der Lage sind soziale Normen zu verstehen, jedoch die emotionale Antwort auf eine Normverletzung beeinträchtigt ist. Die gefundenen

elektrophysiologischen Defizite könnten daher die gut belegten emotionalen Defizite bei antisozialen Individuen widerspiegeln (Dawel et al., 2012; Hoppenbrouwers et al., 2016; Marsh & Blair, 2008), da die MFN vor allem eine affektive Ergebnisevaluation darstellt (Boksem et al., 2006; Gehring & Willoughby, 2002; Masaki et al., 2006). Zuletzt könnte die Attenuierung aus grundsätzlichen neurophysiologischen Beeinträchtigungen entstanden sein, welche unabhängig vom Normverständnis oder spezifischen emotionalen Reaktionen sind. Wie bereits erwähnt könnten hier Veränderungen im anterioren cingulären Cortex eine Rolle spielen (Glenn et al., 2009; Glenn et al., 2010; Kiehl, 2006; Kiehl et al., 2001; Koenigs, 2012; Ly et al., 2012; Rilling et al., 2007; Yang & Raine, 2009).

3.3 Einfluss von Emotionen und Empathie bei antisozialen Individuen

Insgesamt deutet die Studienlage auf äquivalente sozio-kontextuelle Determinanten bei Entscheidungsprozessen in antisozialen und gesunden Populationen hin. Sowohl die Gruppenzugehörigkeit wie auch die Erwartung mehrfacher Interaktionen wirkten sich auf dieselbe Art und Weise bei beiden Gruppen aus. Vor dem Hintergrund neurophysiologischer Veränderungen bei antisozialen Individuen wurde die Hypothese formuliert, dass diese Befunde emotionale Defizite widerspiegeln. Diese Überlegung deckt sich mit der Idee und den Befunden, dass Defizite bezüglich empathischer Fähigkeiten zu aggressivem Verhalten führen (Jolliffe & Farrington, 2004; Miller & Eisenberg, 1988).

Es stellt sich daher die weitere Frage, ob Entscheidungsprozesse und das Verhalten antisozialer Individuen durch emotionale Faktoren beeinflusst werden können. Studie 1 zeigte bei Gesunden deutlich prosozialeres Verhalten nach einer Empathie-Induktion. Sollten emotionale Prozesse bei antisozialen Individuen von besonderer Relevanz sein, so sollten in diesem Paradigma deutliche Unterschiede im Vergleich zu gesunden Kontrollen auftreten. Mayer, Jusyte, et al. (2018) führten dasselbe Paradigma aus Studie 1 bei inhaftierten Gewaltstraftätern durch. Zudem wurden verschiedene subjektive und objektive Maße zur Erfassung empathischer Fähigkeiten erhoben. Überraschenderweise konnte diese Studie keine Defizite bezüglich empathischer Fähigkeiten bei inhaftierten Gewaltstraftätern finden. Die Empathie-Induktion führte zudem wie bei gesunden Kontrollen zu einer erhöhten Bereitschaft für prosoziales Verhalten. Diese Daten

bestärken die Annahme, dass antisoziale Individuen in ökonomischen Paradigmen und den damit verbundenen sozialen Entscheidungsprozessen keine Defizite aufweisen. Dennoch konnte diese Untersuchung Zusammenhänge zwischen hoher Psychopathie und weniger selbstberichteter Empathie, ein reduziertes affektives Ansprechen auf die Empathie-Induktion sowie weniger ausgeprägtes altruistisches Verhalten nachweisen. Somit wiederholt sich das vorige Fazit, dass nicht Antisozialität per se, jedoch spezifische Subtypen, mit veränderten Entscheidungsprozessen einhergehen.

Eine weitere Studie von Klapwijk et al. (2016) untersuchte Jugendliche mit einer Störung des Sozialverhaltens, die im Justizsystem auffällig geworden waren. Diese Jugendlichen erhielten emotionales Feedback zu ihren Entscheidungen in einer Variante des Dictator Games (wütend, enttäuscht oder zufrieden). Nach diesem Feedback spielten die Teilnehmer weitere Durchgänge des Dictator Games. Hierbei wurden zusätzlich neurale Mechanismen mittels fMRI aufgezeichnet. Die Entscheidungen von Jugendlichen mit einer Störung des Sozialverhaltens änderten sich nicht durch die emotionale Rückmeldung, während gesunde Kontrollen nach einer zufriedenen Rückmeldung unfairere Angebote machten. Die Neuroimaging Daten deuteten wie die Verhaltensdaten laut den Autoren auf Schwierigkeiten bei der expliziten Verarbeitung emotionaler Hinweisreize von anderen hin.

Diese beiden Studien illustrieren erneut, dass trotz theoretisch klarer Herleitung die Befundlage zur sozialen Entscheidungsfindung bei antisozialen Individuen nicht eindeutig ist. Trotz vielbelegter Defizite bei empathischen Fähigkeiten (Jolliffe & Farrington, 2004), scheinen sich diese nicht bei der Entscheidungsfindung auszuwirken (Mayer, Jusyte, et al., 2018). Die Emotionserkennung von Gesichtern bei Gewaltstraftätern ist ebenfalls durch eine Vielzahl von Studien als beeinträchtigt einzuschätzen (Chapman, Gillespie, & Mitchell, 2018), zudem scheint auch die vokale Emotionserkennung bei Psychopathen verändert zu sein (Dawel et al., 2012). Trotz dieser Befunde scheinen sich emotionale Rückmeldungen von Interaktionspartnern nicht negativ bei Menschen aus dem antisozialen Spektrum auszuwirken (Klapwijk et al., 2016). Aus den bisherigen Studien zu emotionalen Einflüssen auf Entscheidungsprozesse bei antisozialen Individuen kann also erneut das Fazit gezogen werden, dass diese Prozesse allem Anschein nach nicht so gestört sind wie

angenommen. Die Rolle von psychopathischen Eigenschaften und deren Facetten muss durch weitere Forschung untersucht werden.

3.4 Kritische Würdigung der verwendeten Paradigmen

Die hier referierten Paradigmen können zu Recht als prominente Standardinstrumente bezeichnet werden, da diese etliche Male in verschiedenen wissenschaftlichen Disziplinen angewandt wurden. Trotz dieser Prominenz lassen sich methodische Probleme bei der Verwendung der Paradigmen illustrieren. Zum Beispiel untersuchten Forsythe et al. (1994) in der erstmaligen Anwendung des Paradigmas, ob Fairness alleine die Bereitschaft etwas in diesem Spiel zu teilen determiniert. Die Autoren kamen in ihrer Studie zum Schluss, dass Fairness alleine das gezeigte Verhalten nicht ausreichend erklären konnte. In anderen Publikationen ist wiederum oft die Rede davon, dass die Daten aus dem Dictator Game unter anderem durch eine Ungleichheitsaversion (Bolton & Ockenfels, 1998; Fehr & Schmidt, 1999) oder Altruismus (Andreoni & Miller, 2002) erklärt werden können. Es scheint also zunächst nicht endgültig geklärt zu sein, welches Konstrukt mit dem Dictator Game erfasst wird.

Hinzu kommt, dass andere Autoren mit experimentellen Untersuchungen die Validität des Dictator Games infrage stellten. List (2007) und Bardsley (2008) argumentieren, dass das Verhalten im Dictator Game zu einem großen Anteil ein Artefakt der experimentellen Situation ist. Dafür veränderten die Autoren in ihren Studien die Instruktionen des Paradigmas dahingehend, dass es hier zusätzlich für den Diktator die Option gab, Geld von seinem Mitspieler zu nehmen. Dadurch gaben Versuchspersonen wesentlich weniger Geld ab, viele Probanden nahmen sogar nur Geld von ihren Mitspielern. Fairness, Ungleichheit oder Altruismus schienen in diesen Experimenten eine nur noch untergeordnete Rolle zu spielen. Es scheint also von zentraler Bedeutung zu sein, wie die Aufgabe formuliert wird beziehungsweise welche implizite Aufforderung in den Instruktionen transportiert wird. Die Experimente von List (2007) und Bardsley (2008) lassen den Schluss zu, dass letztendlich in diesem Paradigma nur eine Antwortverzerrung (so genannter response bias) abgebildet wird, da eine „wahre“ Entscheidung nicht mit der Anzahl der Optionen variieren sollte. Eine weitere Studie

bestätigt die Kritik am Dictator Game und dass hier ein Effekt der Antwortalternativen vorliegt (Cappelen, Nielsen, Sørensen, Tungodden, & Tyran, 2013). Die Autoren dieser Studie schlussfolgern zwei mögliche Ansätze zu den gefundenen Ergebnissen. Das Verhalten im Dictator Game könnte demnach zumindest zu einem Teil von dem Wunsch motiviert sein, entweder zu signalisieren, dass man nicht ausschließlich eigennützig ist (Andreoni & Bernheim, 2009) oder durch den Wunsch einer sozialen Norm nach zu handeln, welche wiederum von den Antwortalternativen abhängig ist (Bardsley, 2008; List, 2007). In jedem Fall kann hieraus das Zwischenfazit gezogen werden, dass zumindest eine kritische Haltung den Ergebnissen aus Untersuchungen mit einem solchen Paradigma gerechtfertigt ist. Eine Untersuchung des Einflusses der Antwortalternativen im Ultimatum Game oder Give Some Dilemma wurde zum jetzigen Zeitpunkt noch nicht durchgeführt. Durch die Ähnlichkeit der Paradigmen könnten jedoch ähnliche Resultate zu erwarten sein.

Zudem könnten bei den genannten Paradigmen durch den Faktor der Anonymität beeinflusst werden. Es könnte sein, dass die Teilnehmer eines eigentlich „anonymen“ Experiments sich aufgrund des Kontextes (Aufnahme der Antworten des Teilnehmers, Anwesenheit des Versuchsleiters) verhalten als würden sie beobachtet werden (Gigerenzer & Gigerenzer, 2005; Haley & Fessler, 2005). Es konnte in experimentellen Studien belegt werden, dass Probanden sehr sensitiv gegenüber der Möglichkeit, beobachtet zu werden, sind. In einer eigenen Studie sowie einer Meta-Analyse der bisherigen Literatur konnten Nettle et al. (2013) zeigen, dass bereits sehr subtile Hinweise, wie ein Bild von beobachtenden Augen oder sogar nur ein aus Punkten stilisiertes Gesicht, die Häufigkeit von prosozialem Verhalten im Dictator Game erhöhten. Es ist anzumerken, dass dennoch nicht die durchschnittliche Anzahl von Punktabgaben verändert wurden. Insgesamt lassen sich Resultate zu den sogenannten „watching eyes“ (zu Deutsch, die beobachtenden Augen) nicht eindeutig einordnen. Es bleibt jedoch auch anzumerken, dass diese Befunde einen weiteren Hinweis auf eine mögliche Konfundierung der Studien zu ökonomischen Spielen darstellen. Gerade beim Vergleich von gesunden Personen zu inhaftierten Menschen könnte die Anwesenheit eines Versuchsleiters zu unterschiedlichen Effekten und damit Verzerrungen der Ergebnisse führen.

Ein abschließendes Problem bei der Bewertung der ökonomischen Spiele ist die externe Validität. Neben den Aspekten der Anonymität und den dargebotenen Antwortalternativen, scheint es bisher ungeklärt, ob die erfassten Konstrukte im Labor Zusammenhänge mit dem Verhalten im „echten“ Leben aufweisen. Zwar werden den ökonomischen Spielen gute psychometrische Eigenschaften in Form von Reliabilität und Stabilität zugeschrieben, welche sich eignen um objektiv (und unabhängig von einer Selbsteinschätzung der Versuchsperson) Persönlichkeitseigenschaften zu testen (Baumert, Schlösser, & Schmitt, 2014). Des Weiteren finden sich in der Literatur Belege für, aber auch gegen die externe Validität der Paradigmen. Einerseits wurden keine Zusammenhänge zwischen den Abgaben in einem Dictator Game im Labor und dem Aufteilen eines Geldbetrages in einem natürlichen Kontext zwischen den Teilnehmern (ohne dessen Wissen, an einer Studie teilzunehmen) und einem Fremden gefunden (Winking & Mizer, 2013). Andererseits fand eine andere Studie, dass das Verhalten im Dictator Game das Zurücksenden eines falsch adressierten Briefs (mit Geld darin) vorhersagte und somit diesem Paradigma eine gute externe Validität bescheinigt werden konnte (Franzen & Pointner, 2013). In einer neueren systematischen Untersuchung der externalen Validität von ökonomischen Spielen, die zusätzlich eine Literaturübersicht sowie eine Meta-Analyse einschließt, wurden per Selbsteinschätzung soziale Verhaltensweisen in der Vergangenheit, Entscheidungen in sieben ökonomischen Spielen (unter anderem dem Dictator Game und dem Ultimatum Game) sowie fünf naturalistische Feldexperimente erhoben (Galizzi & Navarro-Martinez, 2019). Es konnte gezeigt werden, dass ökonomische Spiele soziales Verhalten im Feld sowie vergangenes soziales Verhalten nicht gut erklären können, am eindrücklichsten war, dass es überhaupt keinen Zusammenhang zwischen dem Verhalten im Labor und in einer natürlichen Umgebung gab. In der Übersicht und der Meta-Analyse zur bisherigen Literatur stellten die Autoren verschiedene methodische Probleme vergangener Studien zu diesem Thema fest, unter anderem wurde nur ein Spiel mit einer (oft ähnlichen) Feldvariable verglichen und zudem gab es eine große Streuung bezüglich der verwendeten Maße, Methodik und Stichproben. Außerdem fehlte es an einer klaren theoretischen Idee bezüglich der Übersetzung von ökonomischen Spielen in die „echte Welt“ beziehungsweise, in welcher Art und Weise die verschiedenen erhobenen Variablen mit den Spielen zusammenhängen könnten. Dadurch entsteht laut den

Autoren die Gefahr, dass (vor allem vor dem Hintergrund, dass vor allem positive Ergebnisse veröffentlicht werden, so genannter Publikations-Bias), bisher gefundene Ergebnisse aus fragwürdigen Korrelationen abgeleitet wurden. Unabhängig davon finden Galizzi and Navarro-Martinez (2019), dass weniger als 40% der betrachteten Studien Zusammenhänge zwischen dem Verhalten im Labor und im Feld finden. Insgesamt sollten Studienergebnisse mit ökonomischen Spielen kritisch bezüglich der referierten Aspekte diskutiert werden. Zudem sind weitere Forschungsarbeiten notwendig, um diese Limitationen besser zu verstehen und neuere experimentelle Designs zu finden, welche diese Kritikpunkte verbessern.

3.5 Fazit

Diese Dissertation beschäftigte sich mit sozialen Entscheidungsprozessen und intrapsychischen sowie sozio-kontextuellen Determinanten dieses Verhaltens. In drei Studien wurden verschiedene Aspekte dieser Prozesse vor dem Hintergrund unterschiedlicher sozio-kontextueller und individueller Charakteristika untersucht. Bei gesunden Individuen konnten zahlreiche Befunde vergangener Studien repliziert und erweitert werden. Gerade die Verbindung von ökonomischen Paradigmen und dem psychologischen Konstrukt der Empathie sowie die Anwendung eines Minimal-Group Paradigmas im Give Some Dilemma sind hervorzuheben. Zudem wurde in zwei Studien eine extrem relevante und schwer betroffene Stichprobe aus inhaftierten Gewaltstraftätern rekrutiert – ein Kollektiv zu dem es trotz massiver individueller und gesellschaftlicher Kosten (Scott et al., 2001) bisher zu wenig Forschungsarbeit gibt. Bezüglich Studie 2 und Studie 3 ist zu betonen, dass vor allem die Übertragung der sozio-kontextuellen Faktoren sowie die Erfassung elektrophysiologischer Korrelate ein wichtiger Schritt zur Erforschung der antisozialen Persönlichkeitsstörung darstellen. Eine Vielzahl von Experimenten und die Daten aus dieser Dissertation liefern keine Hinweise auf ein verändertes oder rationaleres Verhalten in ökonomischen Spielen (im Vergleich zu gesunden Individuen). Diese Erkenntnisse sollten durch zukünftige Studien ergänzt werden und könnten so eine wichtige Rolle bei der Behandlung dieser Erkrankung spielen.

Literatur

- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC.
- Andreoni, J., & Bernheim, B. D. (2009). Social image and the 50–50 norm: A theoretical and experimental analysis of audience effects. *Econometrica*, *77*(5), 1607-1636. doi:10.3982/ECTA7384
- Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *70*(2), 737-753. doi:10.1111/1468-0262.00302
- Arbuckle, N. L., & Cunningham, W. A. (2012). Understanding everyday psychopathy: shared group identity leads to increased concern for others among undergraduates higher in psychopathy. *Social Cognition*, *30*(5), 564-583. doi:10.1521/soco.2012.30.5.564
- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390-1396. doi:0.1126/science.7466396
- Balliet, D., Wu, J., & De Dreu, C. K. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, *140*(6), 1556-1581. doi:10.1037/a0037737
- Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics*, *11*(2), 122-133. doi:10.1007/s10683-007-9172-2
- Batson, C. D. (1990). How social an animal? The human capacity for caring. *American Psychologist*, *45*(3), 336. doi:10.1037/0003-066X.45.3.336
- Baumert, A., Schlösser, T., & Schmitt, M. (2014). Economic Games: A performance-based assessment of fairness and altruism. *European Journal of Psychological Assessment*, *30*(3), 178-192. doi:10.1027/1015-5759/a000183
- Berg, J. M., Lilienfeld, S. O., & Waldman, I. D. (2013). Bargaining with the devil: Using economic decision-making tasks to examine the heterogeneity of psychopathic traits. *Journal of Research in Personality*, *47*(5), 472-482. doi:10.1016/j.jrp.2013.04.003
- Black, D. W., Baumgard, C. H., Bell, S. E., & Kao, C. (1996). Death rates in 71 men with antisocial personality disorder: A comparison with general population mortality. *Psychosomatics*, *37*(2), 131-136. doi:10.1016/S0033-3182(96)71579-7
- Boksem, M. A. S., & De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Social Neuroscience*, *5*(1), 118-128. doi:10.1080/17470910903202666
- Boksem, M. A. S., Tops, M., Wester, A. E., Meijman, T. F., & Lorist, M. M. (2006). Error-related ERP components and individual differences in punishment and reward sensitivity. *Brain Res*, *1101*(1), 92-101. doi:10.1016/j.brainres.2006.05.004
- Bolton, G. E., & Ockenfels, A. (1998). Strategy and Equity: An ERC-Analysis of the Güth–van Damme Game. *Journal of Mathematical Psychology*, *42*(2), 215-226. doi:10.1006/jmps.1998.1211
- Bolton, G. E., & Zwick, R. (1995). Anonymity versus punishment in ultimatum bargaining. *Games and Economic Behavior*, *10*(1), 95-121. doi:10.1006/game.1995.1026
- Brandstätter, H., & Königstein, M. (2001). Personality influences on ultimatum bargaining decisions. *European Journal of Personality*, *15*(S1), S53-S70. doi:10.1002/per.424

- Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, 86(2), 307-324. doi:10.1037/0033-2909.86.2.307
- Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of Social Issues*, 55(3), 429-444. doi:10.1111/0022-4537.00126
- Burton-Chellew, M. N., Nax, H. H., & West, S. A. (2015). Payoff-based learning explains the decline in cooperation in public goods games. *Proceedings of the Royal Society B: Biological Sciences*, 282(1801), 20142678. doi:10.1098/rspb.2014.2678
- Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, 4(6), 215-222. doi:10.1016/S1364-6613(00)01483-2
- Camerer, C. F. (2003). Strategizing in the brain. *Science*, 300(5626), 1673-1675. doi:10.1126/science.1086215
- Cappelen, A. W., Nielsen, U. H., Sørensen, E. Ø., Tungodden, B., & Tyran, J.-R. (2013). Give and take in dictator games. *Economics Letters*, 118(2), 280-283. doi:10.1016/j.econlet.2012.10.030
- Chapman, H., Gillespie, S. M., & Mitchell, I. J. (2018). Facial affect processing in incarcerated violent males: A systematic review. *Aggression and Violent Behavior*, 38, 123-138. doi:10.1016/j.avb.2017.10.006
- Chmura, T., Engel, C., & Englerth, M. (2017). At the mercy of a prisoner three dictator experiments. *Applied Economics Letters*, 24(11), 774-778. doi:10.1080/13504851.2016.1226486
- Coid, J., & Ullrich, S. (2010). Antisocial personality disorder is on a continuum with psychopathy. *Comprehensive Psychiatry*, 51(4), 426-433. doi.org/10.1016/j.comppsy.2009.09.006
- Coid, J., Yang, M., Tyrer, P., Roberts, A., & Ullrich, S. (2006). Prevalence and correlates of personality disorder in Great Britain. *British Journal of Psychiatry*, 188(5), 423-431. doi:10.1192/bjp.188.5.423
- Coid, J., Yang, M., Ullrich, S., Roberts, A., & Hare, R. D. (2009). Prevalence and correlates of psychopathic traits in the household population of Great Britain. *International Journal of Law and Psychiatry*, 32(2), 65-73. doi:10.1016/j.ijlp.2009.01.002
- Compton, W. M., Conway, K. P., Stinson, F. S., Colliver, J. D., & Grant, B. F. (2005). Prevalence, Correlates, and Comorbidity of DSM-IV Antisocial Personality Syndromes and Alcohol and Specific Drug Use Disorders in the United States: Results From the National Epidemiologic Survey on Alcohol and Related Conditions. *The Journal of Clinical Psychiatry*, 66(6), 677-685. doi:10.4088/JCP.v66n0602
- Curry, O., Chesters, M. J., & Viding, E. (2011). The psychopath's dilemma: The effects of psychopathic personality traits in one-shot games. *Personality and Individual Differences*, 50(6), 804-809. doi:10.1016/j.paid.2010.12.036
- Dawel, A., O'Kearney, R., McKone, E., & Palermo, R. (2012). Not just fear and sadness: Meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neuroscience & Biobehavioral Reviews*, 36(10), 2288-2304. doi:10.1016/j.neubiorev.2012.08.006
- de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *59(1)*, 279-300. doi:10.1146/annurev.psych.59.103006.093625

- Eckel, C. C., & Grossman, P. J. (1996). Altruism in anonymous dictator games. *Games and Economic Behavior*, 16(2), 181-191. doi:10.1006/game.1996.0081
- Edele, A., Dziobek, I., & Keller, M. (2013). Explaining altruistic sharing in the dictator game: The role of affective empathy, cognitive empathy, and justice sensitivity. *Learning and Individual Differences*, 24, 96-102. doi:10.1016/j.lindif.2012.12.020
- Eisenberg, N., & Fabes, R. A. (1990). Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion*, 14(2), 131-149. doi:10.1007/BF00991640
- Eisenberg, N., & Miller, P. A. (1987). The relation of empathy to prosocial and related behaviors. *Psychological Bulletin*, 101(1), 91-119. doi:10.1037/0033-2909.101.1.91
- Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, 302(5643), 290-292. doi:10.1126/science.1089134
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, 14(4), 583-610. doi:10.1007/s10683-011-9283-7
- Fazel, S., & Danesh, J. (2002). Serious mental disorder in 23 000 prisoners: A systematic review of 62 surveys. *The Lancet*, 359(9306), 545-550. doi:10.1016/S0140-6736(02)07740-1
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 817-868. doi:10.1162/003355399556151
- Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3), 347-369. doi:10.1006/game.1994.1021
- Franzen, A., & Pointner, S. (2013). The external validity of giving in the dictator game. *Experimental Economics*, 16(2), 155-169. doi:10.1007/s10683-012-9337-5
- Furnham, A., Richards, S. C., & Paulhus, D. L. (2013). The dark triad of personality: A 10 year review. *Social and Personality Psychology Compass*, 7(3), 199-216. doi:10.1111/spc3.12018
- Galizzi, M. M., & Navarro-Martinez, D. (2019). On the external validity of social preference games: A systematic lab-field study. *Management Science*, 65(3), 976-1002. doi:10.1287/mnsc.2017.2908
- Gehring, W. J., & Willoughby, A. R. (2002). The medial frontal cortex and the rapid processing of monetary gains and losses. *Science*, 295(5563), 2279-2282. doi:10.1126/science.1066893
- Gervais, M. M., Kline, M., Ludmer, M., George, R., & Manson, J. H. (2013). The strategy of psychopathy: primary psychopathic traits predict defection on low-value relationships. *Proceedings of the Royal Society of London B: Biological Sciences*, 280(1757), 20122773. doi:10.1098/rspb.2012.2773
- Gibbon, S., Duggan, C., Stoffers, J., Huband, N., Völlm, B. A., Ferriter, M., & Lieb, K. (2010). Psychological interventions for antisocial personality disorder. *Cochrane Database of Systematic Reviews*(6). doi:10.1002/14651858.CD007668.pub2
- Gigerenzer, G., & Gigerenzer, T. (2005). Is the Ultimatum Game a three-body affair? *Behavioral and Brain Sciences*, 28(6), 823-824. doi:10.1017/S0140525X05300148
- Gillespie, S. M., Mitchell, I. J., Johnson, I., Dawson, E., & Beech, A. R. (2013). Exaggerated intergroup bias in economical decision making games: differential

- effects of primary and secondary psychopathic traits. *PloS One*, 8(8), e69565. doi:10.1371/journal.pone.0069565
- Gintis, H. (2000). Beyond Homo economicus: evidence from experimental economics. *Ecological Economics*, 35(3), 311-322. doi:10.1016/S0921-8009(00)00216-0
- Glenn, A. L., Raine, A., & Schug, R. A. (2009). The neural correlates of moral decision-making in psychopathy. *Molecular Psychiatry*, 14(1), 5-6. doi:10.1038/mp.2008.104
- Glenn, A. L., Yang, Y., Raine, A., & Colletti, P. (2010). No volumetric differences in the anterior cingulate of psychopathic individuals. *Psychiatry Research: Neuroimaging*, 183(2), 140-143. doi:10.1016/j.psychresns.2010.05.009
- Goodwin, R. D., & Hamilton, S. P. (2003). Lifetime comorbidity of antisocial personality disorder and anxiety disorders among adults in the community. *Psychiatry Research*, 117(2), 159-166. doi:10.1016/S0165-1781(02)00320-7
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4), 367-388. doi:10.1016/0167-2681(82)90011-7
- Haley, K. J., & Fessler, D. M. T. (2005). Nobody's watching?: Subtle cues affect generosity in an anonymous economic game. *Evolution and Human Behavior*, 26(3), 245-256. doi:10.1016/j.evolhumbehav.2005.01.002
- Hare, R. D. (2003). *The Hare Psychopathy Checklist Revised (2nd ed.)*. Toronto, Canada: Multi-Health Systems.
- Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M. G. H., Holroyd, C. B., & Miltner, W. H. R. (2011). Why humans deviate from rational choice. 48(4), 507-514. doi:10.1111/j.1469-8986.2010.01081.x
- Hilbig, B. E., Thielmann, I., Hepp, J., Klein, S. A., & Zettler, I. (2015). From personality to altruistic behavior (and back): Evidence from a double-blind dictator game. *Journal of Research in Personality*, 55, 46-50. doi:10.1016/j.jrp.2014.12.004
- Hoppenbrouwers, S. S., Bulten, B. H., & Brazil, I. A. (2016). Parsing fear: A reassessment of the evidence for fear deficits in psychopathy. *Psychological Bulletin*, 142(6), 573-600. doi:10.1037/bul0000040
- Johnston, L., Hawes, D. J., & Straiton, M. (2014). Psychopathic Traits and Social Cooperation in the Context of Emotional Feedback. *Psychiatry, Psychology and Law*, 21(5), 767-778. doi:10.1080/13218719.2014.893550
- Jolliffe, D., & Farrington, D. P. (2004). Empathy and offending: A systematic review and meta-analysis. *Aggression and Violent Behavior*, 9(5), 441-476. doi:10.1016/j.avb.2003.03.001
- Khadjavi, M., & Lange, A. (2013). Prisoners and their dilemma. *Journal of Economic Behavior & Organization*, 92, 163-175. doi:10.1016/j.jebo.2013.05.015
- Kiehl, K. A. (2006). A cognitive neuroscience perspective on psychopathy: evidence for paralimbic system dysfunction. *Psychiatry Research*, 142(2), 107-128. doi:10.1016/j.psychres.2005.09.013
- Kiehl, K. A., Smith, A. M., Hare, R. D., Mendrek, A., Forster, B. B., Brink, J., & Liddle, P. F. (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Biological Psychiatry*, 50(9), 677-684. doi:10.1016/S0006-3223(01)01222-7
- Klapwijk, E. T., Lelieveld, G.-J., Aghajani, M., Boon, A. E., van der Wee, N. J., Popma, A., . . . Colins, O. F. (2016). Fairness decisions in response to emotions: a functional MRI study among criminal justice-involved boys with conduct disorder.

- Social Cognitive and Affective Neuroscience*, 11(4), 674-682.
doi:10.1093/scan/nsv150
- Koenigs, M. (2012). The role of prefrontal cortex in psychopathy. *Reviews in the Neurosciences*, 23(3), 253-262. doi:10.1515/revneuro-2012-0036
- Koenigs, M., Kruepke, M., & Newman, J. P. (2010). Economic decision-making in psychopathy: a comparison with ventromedial prefrontal lesion patients. *Neuropsychologia*, 48(7), 2198-2204.
doi:10.1016/j.neuropsychologia.2010.04.012
- Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual review of sociology*, 24(1), 183-214. doi:10.1146/annurev.soc.24.1.183
- Lenzenweger, M. F., Lane, M. C., Loranger, A. W., & Kessler, R. C. (2007). DSM-IV personality disorders in the national comorbidity survey replication. *Biological Psychiatry*, 62(6), 553-564. doi:10.1016/j.biopsych.2006.09.019
- List, John A. (2007). On the interpretation of giving in dictator games. *Journal of Political Economy*, 115(3), 482-493. doi:10.1086/519249
- Ly, M., Motzkin, J. C., Philippi, C. L., Kirk, G. R., Newman, J. P., Kiehl, K. A., & Koenigs, M. (2012). Cortical thinning in psychopathy. *American Journal of Psychiatry*, 169(7), 743-749. doi:10.1176/appi.ajp.2012.11111627
- Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, 32(3), 454-465. doi:10.1016/j.neubiorev.2007.08.003
- Masaki, H., Takeuchi, S., Gehring, W. J., Takasawa, N., & Yamazaki, K. (2006). Affective-motivational influences on feedback-related ERPs in a gambling task. *Brain Res*, 1105(1), 110-121. doi:10.1016/j.brainres.2006.01.022
- Mayer, S. V., Jusyte, A., Klimecki-Lenz, O. M., & Schönberg, M. (2018). Empathy and altruistic behavior in antisocial violent offenders with psychopathic traits. *Psychiatry Research*, 269, 625-632. doi:10.1016/j.psychres.2018.08.035
- Mayer, S. V., Rauss, K., Pourtois, G., Jusyte, A., & Schönberg, M. (2018). Behavioral and electrophysiological responses to fairness norm violations in antisocial offenders. *European Archives of Psychiatry and Clinical Neuroscience*.
doi:10.1007/s00406-018-0878-2
- Miller, P. A., & Eisenberg, N. (1988). The relation of empathy to aggressive and externalizing/antisocial behavior. *Psychological Bulletin*, 103(3), 324-344.
doi:10.1037/0033-2909.103.3.324
- Mokros, A., Menner, B., Eisenbarth, H., Alpers, G. W., Lange, K. W., & Osterheider, M. (2008). Diminished cooperativeness of psychopaths in a prisoner's dilemma game yields higher rewards. *Journal of Abnormal Psychology*, 117(2), 406-413.
doi:10.1037/0021-843X.117.2.406
- Nettle, D., Harper, Z., Kidson, A., Stone, R., Penton-Voak, I. S., & Bateson, M. (2013). The watching eyes effect in the Dictator Game: it's not how much you give, it's being seen to give something. *Evolution and Human Behavior*, 34(1), 35-40.
doi:10.1016/j.evolhumbehav.2012.08.004
- Ortmann, A., & Tichy, L. K. (1999). Gender differences in the laboratory: evidence from prisoner's dilemma games. *Journal of Economic Behavior & Organization*, 39(3), 327-339. doi:10.1016/S0167-2681(99)00038-4
- Osumi, T., & Ohira, H. (2010). The positive side of psychopathy: Emotional detachment in psychopathy and rational decision-making in the ultimatum game. *Personality and Individual Differences*, 49(5), 451-456. doi:10.1016/j.paid.2010.04.016

- Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, 36(6), 556-563. doi:10.1016/S0092-6566(02)00505-6
- Persky, J. (1995). The Ethology of Homo Economicus. *Journal of Economic Perspectives*, 9(2), 221-231. doi:10.1257/jep.9.2.221
- Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, 48(5), 656-664. doi:10.1111/j.1469-8986.2010.01136.x
- Polezzi, D., Daum, I., Rubaltelli, E., Lotto, L., Civai, C., Sartori, G., & Rumiati, R. (2008). Mentalizing in economic decision-making. *Behavioural Brain Research*, 190(2), 218-223. doi:10.1016/j.bbr.2008.03.003
- Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When things are better or worse than expected: the medial frontal cortex and the allocation of processing resources. *J Cogn Neurosci*, 18(7), 1112-1119. doi:10.1162/jocn.2006.18.7.1112
- Radke, S., Brazil, I. A., Scheper, I., Bulten, B. H., & De Bruijn, E. R. (2013). Unfair offers, unfair offenders? Fairness considerations in incarcerated individuals with and without psychopathy. doi:10.3389/fnhum.2013.00406
- Rapoport, A. (1989). Prisoner's Dilemma. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *Game Theory* (pp. 199-204). London: Palgrave Macmillan UK.
- Rapoport, A., & Chammah, A. M. (1965). *Prisoner's dilemma: A study in conflict and cooperation* (Vol. 165): University of Michigan press.
- Rilling, J. K., Glenn, A. L., Jairam, M. R., Pagnoni, G., Goldsmith, D. R., Elfenbein, H. A., & Lilienfeld, S. O. (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biological Psychiatry*, 61(11), 1260-1271. doi:10.1016/j.biopsych.2006.07.021
- Robins, L. N., & Price, R. K. (1991). Adult disorders predicted by childhood conduct problems: Results from the NIMH Epidemiologic Catchment Area project. *Psychiatry: Interpersonal and Biological Processes*, 54(2), 116-132. doi:10.1080/00332747.1991.11024540
- Sally, D. (1995). Conversation and cooperation in social dilemmas a meta-analysis of experiments from 1958 to 1992. *Rationality and Society*, 7(1), 58-92. doi:10.1177/1043463195007001004
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, 300(5626), 1755-1758. doi:10.1126/science.1082976
- Saslow, L. R., Willer, R., Feinberg, M., Piff, P. K., Clark, K., Keltner, D., & Saturn, S. R. (2013). My Brother's Keeper?: Compassion Predicts Generosity More Among Less Religious Individuals. *Social Psychological and Personality Science*, 4(1), 31-38. doi:10.1177/1948550612444137
- Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport*, 16(4), 407-411. doi:10.1097/00001756-200503150-00020
- Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: follow up study of antisocial children into adulthood. *British Medical Journal*, 323(7306), 1-5. doi:10.1136/bmj.323.7306.191
- Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the

- cingulate cortex. *Nature Reviews Neuroscience*, 12(3), 154-167.
doi:10.1038/nrn2994
- Small, D. A., & Loewenstein, G. (2003). Helping a victim or helping the victim: Altruism and identifiability. *Journal of Risk and Uncertainty*, 26(1), 5-16.
doi:10.1023/A:1022299422219
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, 1(2), 149-178.
doi:10.1002/ejsp.2420010202
- Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Monterey, CA: Brooks/Cole.
- Telle, N.-T., & Pfister, H.-R. (2012). Not only the miserable receive help: Empathy promotes prosocial behaviour toward the happy. *Current Psychology*, 31(4), 393-413. doi:10.1007/s12144-012-9157-y
- Torgersen, S., Kringlen, E., & Cramer, V. (2001). The prevalence of personality disorders in a community sample. *JAMA Psychiatry*, 58(6), 590-596.
doi:10.1001/archpsyc.58.6.590
- Van Lange, P. A., Klapwijk, A., & Van Munster, L. M. (2011). How the shadow of the future might promote cooperation. *Group Processes & Intergroup Relations*, 14(6), 857-870. doi:10.1177/1368430211402102
- Van Lange, P. A., & Kuhlman, D. M. (1994). Social value orientations and impressions of partner's honesty and intelligence: A test of the might versus morality effect. *Journal of Personality and Social Psychology*, 67(1), 126-141. doi:10.1037/0022-3514.67.1.126
- Van Lange, P. A., Ouwerkerk, J. W., & Tazelaar, M. J. (2002). How to overcome the detrimental effects of noise in social interaction: the benefits of generosity. *Journal of Personality and Social Psychology*, 82(5), 768-780. doi:10.1037/0022-3514.82.5.768
- Vieira, J. B., Almeida, P. R., Ferreira-Santos, F., Barbosa, F., Marques-Teixeira, J., & Marsh, A. A. (2014). Distinct neural activation patterns underlie economic decisions in high and low psychopathy scorers. *Social Cognitive and Affective Neuroscience*, 9(8), 1099-1107. doi:10.1093/scan/nst093
- Wai, M., & Tiliopoulos, N. (2012). The affective and cognitive empathic nature of the dark triad of personality. *Personality and Individual Differences*, 52(7), 794-799. doi:10.1016/j.paid.2012.01.008
- Wang, Y., Zhang, Z., Bai, L., Lin, C., Osinsky, R., & Hewig, J. (2017). Ingroup/outgroup membership modulates fairness consideration: neural signatures from ERPs and EEG oscillations. *Scientific Reports*, 7, 39827. doi:10.1038/srep39827
- Wilson, W. (1971). Reciprocation and other techniques for inducing cooperation in the Prisoner's Dilemma game. *Journal of Conflict Resolution*, 15(2), 167-195.
doi:10.1177/002200277101500205
- Winking, J., & Mizer, N. (2013). Natural-field dictator game shows no altruistic giving. *Evolution and Human Behavior*, 34(4), 288-293.
doi:10.1016/j.evolhumbehav.2013.04.002
- Yamagishi, T., Horita, Y., Takagishi, H., Shinada, M., Tanida, S., & Cook, K. S. (2009). The private rejection of unfair offers and emotional commitment. *Proceedings of the National Academy of Sciences*, 106(28), 11520-11523.
doi:10.1073/pnas.0900636106

Yang, Y., & Raine, A. (2009). Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: A meta-analysis. *Psychiatry Research: Neuroimaging*, 174(2), 81-88. doi:10.1016/j.pscychresns.2009.03.012

Appendix A



**Erklärung nach § 5 Abs. 2 Nr. 8 der Promotionsordnung der Math.-Nat. Fakultät
-Anteil an gemeinschaftlichen Veröffentlichungen-
Nur bei kumulativer Dissertation erforderlich!**

**Declaration according to § 5 Abs. 2 No. 8 of the PhD regulations of the Faculty of
Science**

**-Collaborative Publications-
For Cumulative Theses Only!**

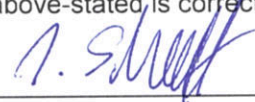
Last Name, First Name: Scheeff, Jonathan

List of Publications

1. Klimecki, O. M., Mayer, S. V. Jusyte, A., Scheeff, J., & Schönberg, M. (2016). "Empathy Promotes Altruistic Behavior in Economic Interactions." *Scientific Reports* 6: 31961. doi:10.1038/srep31961
2. Scheeff, J., Jusyte, A., & Schönberg, M. (2018). Investigating social-contextual determinants of cooperation in incarcerated violent offenders. *Scientific Reports*, 8(1), 17204. doi:10.1038/s41598-018-35450-z
3. Scheeff, J., Rauss, K., & Schönberg, M. (submitted). Effects of group affiliation on neural signatures of fairness norm violations in antisocial violent offenders.

Nr.	Accepted publication yes/no	List of authors	Position of candidate in list of authors	Scientific ideas by the candidate (%)	Data generation by the candidate (%)	Analysis and Interpretation by the candidate (%)	Paper writing done by the candidate (%)
<i>Optionally, you can also declare the above-stated categories in a written statement on a separate sheet of paper.</i>							
1	Yes	5	4	0%	80%	25%	10%
2	Yes	3	1	0%	50%	50%	50%
3	No	3	1	50%	50%	50%	50%

I confirm that the above-stated is correct.

1.8.19. 

Date, Signature of the candidate


I/We certify that the above-stated is correct.


Date, Signature of the doctoral committee or at least of one of the supervisors

PD Dr. Michael Schönberg

Appendix B

SCIENTIFIC REPORTS



OPEN

Empathy promotes altruistic behavior in economic interactions

Olga M. Klimecki^{1,†}, Sarah V. Mayer², Aiste Jusyte³, Jonathan Scheeff² & Michael Schönenberg²

Received: 06 April 2016

Accepted: 01 August 2016

Published: 31 August 2016

What are the determinants of altruism? While economists assume that altruism is mainly driven by fairness norms, social psychologists consider empathy to be a key motivator for altruistic behavior. To unite these two theories, we conducted an experiment in which we compared behavior in a standard economic game that assesses altruism (the so-called Dictator Game) with a Dictator Game in which participants' behavioral choices were preceded either by an empathy induction or by a control condition without empathy induction. The results of this within-subject manipulation show that the empathy induction substantially increased altruistic behavior. Moreover, the increase in experienced empathy predicted over 40% of the increase in sharing behavior. These data extend standard economic theories that altruism is based on fairness considerations, by showing that empathic feelings can be a key motivator for altruistic behavior in economic interactions.

In today's society, we are faced with numerous global challenges, ranging from the spread of dangerous diseases to a major refugee crisis unparalleled in magnitude since World War II¹. If we are to meet these challenges successfully, it is crucial that people engage in altruistic acts such as providing medical support or integrating refugees. Although altruism is defined as a costly act performed for the benefit of another², humans and other animals do indeed exhibit and act upon this genuine concern for another's welfare³. Decades of research in the fields of economics^{2,4} and psychology⁵ have been dedicated to elucidating the determinants of altruistic behavior. However, previous work on altruistic behavior in these two disciplines has largely been carried out in isolation.

In economics, the long held assumption of the homo economicus positing that individuals strive to maximize their gains was abandoned in the 1980s. Empirical evidence has demonstrated that participants behave more generously than predicted by rational choice theory^{6,7}. A substantial part of this work employed fairly simple experimental paradigms, particularly the Dictator Game (DG). In the DG, a participant (the 'dictator') can freely divide a given amount of money between himself and an anonymous recipient, who has no choice but to accept the monetary distribution⁸. While the assumption of the homo economicus predicts that participants will not share any money with the recipient, a meta-analysis taking into account hundreds of studies using the DG has shown that on average, participants give about 30% of their endowment⁴. As this behavior is far more generous than expected, experimental economists have extensively investigated the determinants underlying altruistic behavior. Evidence from these studies has provided the basis for the widely accepted idea that social norms related to fairness (i.e., striving for equal material benefits for oneself and others) play a key role in altruism^{9,10}.

However, the psychological factors that drive this kind of altruistic behavior have been poorly understood. This is surprising considering that psychological research^{5,11,12} has established a close link between altruism and emotional empathy, the capacity to share the feelings of another^{3,12}. In fact, the connection between altruism and empathic feelings for others in need was proposed as early as the 18th century by Adam Smith when he wrote "How selfish soever man may be supposed, there are evidently some principles in his nature, which interest him in the fortune of others, and render their happiness necessary to him, though he derives nothing from it, except the pleasure of seeing it. Of this kind is pity or compassion, the emotion we feel for the misery of others" (ref. 13, page 3). Meta-analytic evidence from psychology corroborates this historical postulate by revealing that empathic states and, to a smaller extent, empathic traits predict altruistic behavior⁵. In line with these findings, the empathy-altruism hypothesis¹¹ has posited that altruistic motivation is elicited by empathy felt for a person in

¹Swiss Centre for Affective Sciences, University of Geneva, Geneva, Switzerland. ²Department of Clinical Psychology and Psychotherapy, University of Tübingen, Tübingen, Germany. ³LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany. [†]Present address: Swiss Centre for Affective Sciences, University of Geneva, Chemin des Mines 9, 1202 Geneva, Switzerland. Correspondence and requests for materials should be addressed to O.M.K. (email: olga.klimecki@unige.ch)

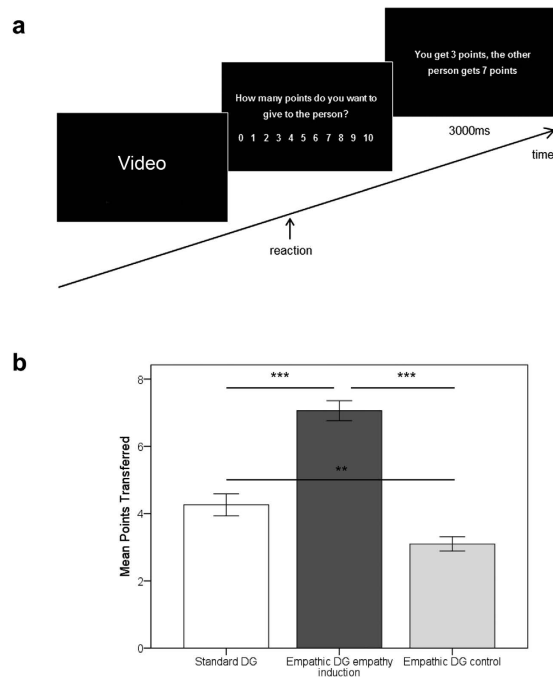


Figure 1. Empathy-induction increased altruistic behavior. (a) Timeline of the Empathic DG. After seeing suffering others (empathy induction) or others in everyday activities (control condition), participants indicated by button press how many points they wanted to share with the person(s) in the video. A subsequent screen presented the results of the decision. (b) Participants ($n = 50$) transferred more points following the empathy induction as opposed to the control condition and the standard DG. Asterisks denote levels of statistical significance of pairwise comparisons with *** $p < 0.001$ and ** $p < 0.01$; error bars denote s.e.m.

need. More recently, researchers have suggested that in both humans and animals empathy has evolved in order to promote altruism towards others in need, pain, or distress³.

Surprisingly, despite these developments in psychology and economics, there have been few attempts to integrate the knowledge from both fields in order to further the understanding of altruism. So far it has been established that empathic traits are positively related to donations in the standard DG¹⁴. Previous research also showed that compassion, which denotes a feeling of care for a suffering other accompanied by the desire to help¹⁵, motivates prosocial behavior in less religious individuals¹⁶. In spite of these first attempts to bridge the psychological concept of empathy (or the related concept of compassion) with economic measures of prosociality, to our knowledge no previous study has tested whether the degree of empathy experienced for the recipient represents the key mechanism motivating prosocial behavior towards this target^{3,5,11} in economic contexts. To address this question, the current study examined whether a within-subject manipulation of empathic states can increase altruistic behavior towards specific others in the DG and whether inter-individual differences in empathic experiences can explain altruistic behavior in this economic context.

Methods

In order to test whether the empathy-altruism hypothesis holds in the realm of economic interactions, we conducted a within-subject experiment in which we compared the standard one-shot DG with a novel version of the DG, the so-called Empathic DG (Fig. 1a). In the Empathic DG, economic choices are either preceded by an empathy induction or by a control condition. The standard one-shot DG, with a recipient who is randomly chosen from a pool of other players, was included to enable a comparison with previous research. Both versions of the DG (standard and Empathic) were played with real monetary stakes (10 monetary units, MUs, per trial). A total of 50 participants (mean age = 23.72 years; 31 females) first played the standard DG, followed by the Empathic DG. In the Empathic DG, empathy was elicited through the presentation of videos depicting a suffering person in need (e.g., a helpless child in an orphanage). This empathy induction was inspired by previous work¹⁷ showing that i) empathy can be induced when participants are confronted with suffering and that ii) the elicited empathic feelings vary between participants. In the control condition of the Empathic DG, participants viewed videos of people performing everyday activities (e.g., two individuals talking to each other). All the videos were selected from the Socio-affective Video Task¹⁸. Gender, race, and age were balanced across empathy-inducing and control videos. This was confirmed by Pearson's chi-square tests that revealed no significant difference for gender ($\chi^2(2) = 2, p = 0.42$), age ($\chi^2(2) = 3.3, p = 0.22$), or race ($\chi^2(1) = 1.5, p = 0.36$). After each of the 44 randomly presented documentary videos (half of them inducing empathy, the other half being control videos), participants determined how many MUs they gave to the person(s) in the video. Following the Empathic DG, participants watched each video again and provided self-reported ratings of positive affect, negative affect, and empathy (on scales ranging from 0, not at all, to 10, very strong). More specifically, we asked them to answer the following

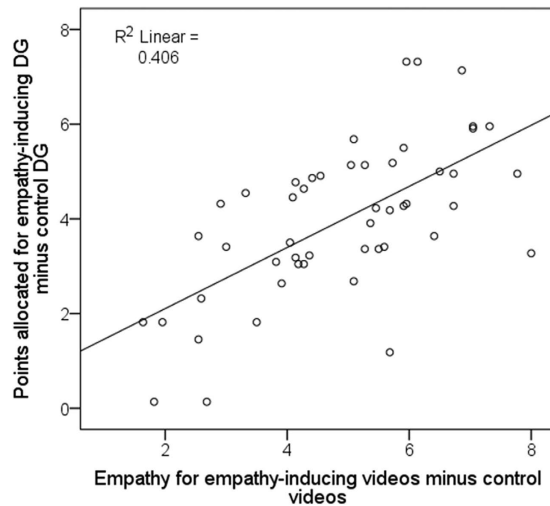


Figure 2. Mean increase of self-reported feelings of empathy from control to empathy-inducing videos predicted mean increase in altruistic behavior from control to empathy-inducing videos ($n = 50$) in the Empathic DG.

questions: “To what degree did you experience positive emotions during the video?”, “To what degree did you experience negative emotions during the video?”, and “How much empathy did you feel?”. Trait empathy was assessed prior to the study using the Interpersonal Reactivity Index (IRI)¹⁹. In line with the definition of fairness as the striving for equal benefits for oneself and others^{9,10}, an even split of MUs between the participant and the recipient was used as an indicator of fairness considerations. All participants provided written informed consent and received monetary compensation for participation. The study protocol was approved by the ethics committee of the University of Tübingen and was carried out in accordance with the approved guidelines and the declaration of Helsinki.

Results

Manipulation check. To check whether empathy-inducing videos elicited different affective states than control videos in the current sample, we first conducted a repeated measures multivariate analysis of variance (MANOVA) with the within subject variable video-type (empathy-inducing vs control videos) and the dependent variables empathy, positive affect, and negative affect. This analysis revealed a significant effect of video-type ($F(3,47) = 236.31, p < 0.001, \eta^2 = 0.94$). Follow-up pairwise comparisons confirmed that empathy-inducing videos indeed elicited more empathy and negative feelings than control videos, while control videos elicited more positive feelings than empathy-inducing videos (all $p < 0.001$; for details, see Supplementary Fig. S1).

Giving behavior. Using a repeated measures ANOVA, we assessed the effect of the empathy induction on altruistic behavior by comparing the amount of MUs shared with the recipient in the standard DG with the amount of MUs shared following the empathy induction or the control condition in the Empathic DG. In line with our assumptions, we observed a large effect of condition ($F(2,48) = 155.35, p < 0.001, \eta^2 = 0.87$), which was significant for all pairwise comparisons (see Fig. 1b for details) and showed that the empathy induction increased helping behavior. Following the empathy induction, participants gave 70.59% of their total endowment, whereas they only gave 42.60% in the standard DG and 30.99% in the control condition.

Regression Analyses. Regression analyses were computed to determine which state components predicted giving behavior. A first regression analysis revealed that empathy significantly predicted generosity in response to empathy-inducing videos ($b = 0.75, p < 0.001$), explaining 31.2% of the variance ($R^2 = 0.31, F(1,48) = 21.73, p < 0.001$; Fig. 2). Negative emotions also predicted generosity in response to empathy-inducing videos ($b = 0.46, p < 0.01$), explaining 19% of the variance ($R^2 = 0.19, F(1,48) = 11.23, p < 0.01$). A multiple regression analysis that simultaneously included both empathy and negative affect revealed that the effect of empathy remained significant ($b = 0.65, p < 0.01$), while the negative affect ratings did not significantly contribute to explaining the variance in this model ($b = 0.11, p = 0.51$).

Subsequently, the effect of the empathy induction on altruistic behavior was examined. To this end, we computed a regression analysis that tested whether the increase in empathy from control videos to empathy-inducing videos predicted the increase in the number of MUs allocated from the control to the empathy-inducing condition. The effect of empathy induction on altruistic behavior was significant ($b = 0.65, p < 0.001$), explaining 40.6% of the variance ($R^2 = 0.41, F(1,48) = 32.79, p < 0.001$; Fig. 2).

Finally, three regression analyses were run to test whether trait empathy (as indicated by the IRI score) predicted altruistic behavior in i) the standard DG, ii) the Empathic DG preceded by an empathy induction and iii) the Empathic DG preceded by a control condition. None of the regressions were significant (all $p \geq 0.36$).

Discussion

The present data show that the empathy-altruism hypothesis^{11,20} can explain behavior in economic interactions. In line with the assumption of fairness norms, behavior in the standard DG and the control condition of the Empathic DG conformed to the prediction of equitable monetary distributions. However, following the empathy induction, participants were willing to give over 70% of their endowments to the suffering others and this increase was explained by an increase in empathic feelings. In other words, it was possible to overcome fairness norms with an empathy induction specifically linked to the suffering recipient. This finding extends the notion that situational empathy is a central motivator of altruism directed at helping others in need, pain, or distress³ to economic interactions. In contrast to previous work, which observed much lower levels of altruism with charities²¹ or fictitious characters²² as recipients, the present study reveals that altruism substantially increases when individuals divide real monetary stakes and when they are presented with visual displays of suffering recipients. The current work showed that the degree of empathy experienced towards a specific person changes as a function of the situation and motivates altruistic behavior in an economic context. This finding extends previous evidence on the relation between trait empathy and altruism towards anonymous strangers¹⁴ as well as the research on the role of compassion as a motivator for prosocial behavior in less religious participants¹⁶. Moreover, our findings provide an explanation for the observation that putting a face to the victims increases altruistic behavior²³ by showing that the extent of empathy experienced towards another person strongly predicts how much people are willing to share in economic interactions.

In the present study, self-reports of empathic feelings predicted a large degree of altruistic behavior, whereas empathic traits were not predictive of altruistic behavior. Although this finding contrasts with previous work, which has reported a significant relationship between trait empathy and behavior in a standard DG¹⁴, it is in line with meta-analytic evidence that prosocial behavior is more strongly related to situational empathy than to empathic traits⁵. These lower correlations for empathic traits might be explained by the heterogeneity of questionnaire measures⁵.

Taken together, the present findings indicate that in order to promote altruism - whether it is for charities, refugees, or in other economic and political contexts - it is essential to appeal to a person's empathy for specific recipients.

References

1. United Nations High Commissioner for Refugees. Global Trends. UNHCR, Geneva, Switzerland (2014).
2. E. Fehr & U. Fischbacher. The nature of human altruism. *Nature* **425**, 785 (2003).
3. F. B. de Waal. Putting the Altruism Back into Altruism: The Evolution of Empathy. *Annual Review of Psychology* **59**, 279 (2007).
4. C. Engel. Dictator games: a meta study. *Experimental Economics* **14**, 583 (2011).
5. N. Eisenberg & P. A. Miller. The relation of empathy to prosocial and related behaviors. *Psychol Bull* **101**, 91 (1987).
6. W. Güth, R. Schmittberger & B. Schwarze. An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization* **3**, 367 (1982).
7. D. Kahneman, J. L. Knetsch & R. H. Thaler. Fairness and the Assumptions of Economics. *J Bus* **59**, S285 (Oct, 1986).
8. R. Forsythe, J. L. Horowitz, N. E. Savin & M. Sefton. Fairness in Simple Bargaining Experiments. *Games and Economic Behavior* **6**, 347 (1994).
9. E. Fehr & K. M. Schmidt. A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics* **114**, 817 (1999).
10. J. Andreoni & B. D. Bernheim. Social image and the 50-50 norm: a theoretical and experimental analysis of audience effects. *Econometrica* **77**, 1607 (2009).
11. C. D. Batson. How social an animal? The human capacity for caring. *American Psychologist* **45**, 336 (1990).
12. N. Eisenberg & R. A. Fabes. Empathy: Conceptualization, measurement, and relation to prosocial behavior. *Motivation and Emotion* **14**, 131 (1990).
13. A. Smith. *The Theory of Moral Sentiments* (Henry G. Bohn, London, 1853). Original work published 1759.
14. A. Edele, I. Dziobek & M. Keller. Explaining altruistic sharing in the dictator game: The role of affective empathy, cognitive empathy, and justice sensitivity. *Learning and individual differences* **24**, 96–102 (2013).
15. J. L. Goetz, D. Keltner & E. Simon-Thomas. Compassion: An Evolutionary Analysis and Empirical Review. *Psychological bulletin* **136**, 351–374 (2010).
16. L. R. Saslow et al. My brother's keeper? Compassion predicts generosity more among less religious individuals. *Social Psychological and Personality Science* **4**, 31–38 (2013).
17. C. D. Batson, K. O'Quin, J. Fultz, M. Vanderplas & A. M. Isen. Influence of Self-Reported Distress and Empathy on Egoistic Versus Altruistic Motivation to Help. *Journal of Personality and Social Psychology* **45**, 706–718 (1983).
18. O. M. Klimecki, S. Leiberg, C. Lamm & T. Singer. Functional Neural Plasticity and Associated Changes in Positive Affect After Compassion Training. *Cerebral Cortex* **23**, 1552 (2013).
19. M. H. Davis. Measuring Individual-Differences in Empathy - Evidence for a Multidimensional Approach. *Journal of Personality and Social Psychology* **44**, 113 (1983).
20. C. D. Batson. *The altruism question: Towards a social-psychological answer* (Erlbaum, Mahwah, NJ, 1991).
21. C. C. Eckel & P. J. Grossman. Altruism in anonymous dictator games. *Games and Economic Behavior* **16**, 181 (1996).
22. N.-T. Telle & H.-R. Pfister. Not only the miserable receive help: empathy promotes prosocial behaviour toward the happy. *Current Psychology* **31**, 393 (2012).
23. D. A. Small & G. Loewenstein. Helping a victim or helping the victim: Altruism and identifiability. *Journal of Risk and Uncertainty* **26**, 5 (2003).

Acknowledgements

We thank Ryan Dutton and Vanessa Sennwald for language editing. Aiste Jusyte is funded by the LEAD Graduate School [GSC1028], a project of the Excellence Initiative of the German federal and state governments.

Author Contributions

O.M.K., S.V.M., A.J. and M.S. designed the experiment, J.S. carried out the experiment, O.M.K., S.V.M. and M.S. analyzed the data and wrote the paper. S.V.M. and O.M.K. prepared Figure 1a, O.M.K. prepared all other figures. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Klimecki, O. M. *et al.* Empathy promotes altruistic behavior in economic interactions. *Sci. Rep.* **6**, 31961; doi: 10.1038/srep31961 (2016).




This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016


Appendix C

SCIENTIFIC REPORTS



OPEN

Investigating social-contextual determinants of cooperation in incarcerated violent offenders

Jonathan Scheeff¹ , Aiste Jusyte² & Michael Schöenberg¹

Antisocial and psychopathic personality traits constitute a severe and treatment resistant form of externalizing psychopathology. While deficits in social information processing have been extensively investigated in these individuals, less is known about their capacity for altruism and cooperation. In particular, no studies to date have investigated whether established social-contextual determinants of cooperation, e.g., group affiliation and number of expected interactions, can motivate cooperative behaviour in antisocial individuals. The present study compared cooperative behaviour of incarcerated violent offenders ($N = 52$) and controls ($N = 46$) by using an established social interaction paradigm (Give Some Dilemma) where two players divide monetary units between themselves and the counterpart. Group affiliation (in- vs. out-group) and number of expected interactions (single-trial vs. repeated-trial interactions) were manipulated. Violent offenders as compared to controls shared less monetary units with their counterparts, indicating an overall reduced cooperation. Both groups showed increased cooperation rates towards in-group members and in repeated interactions. Higher psychopathic traits were associated with lower cooperation in single-trial interactions in the violent offender group. Although cooperation was comparably reduced in violent offenders, behaviour in both groups was determined by the number of expected interactions as well as group affiliation, thus providing evidence for equivalent social-contextual determinants.

Antisocial and psychopathic personality tendencies, such as irritability, aggressiveness, ruthlessness and failure to adhere to social rules and norms, as well as a fundamental lack of care for others, constitute severe and treatment resistant forms of externalizing psychopathology. As a consequence, prosocial behaviours such as altruism or cooperation could be assumed to be uncommon for these individuals. Yet, there is anecdotic and factual evidence suggesting that under certain circumstances, these populations also show a willingness for alliance, e.g., forming criminal enterprises. While deficits in social information processing have been extensively investigated in order to explain aggressive and callous behaviour in these populations^{1–3}, less is known about their capacity for altruism and cooperation. Understanding to what extent prosocial behaviour is impaired and how it can be reinforced is crucial, considering that this may have direct treatment implications for these severe, costly, and treatment-resistant disorders^{4,5}.

Economic games, such as the Dictator Game (DG), the Ultimatum Game (UG), and the Prisoner's Dilemma (PD), have been widely used to study various facets of prosocial behaviour^{6–9}. In the DG, a person ("the dictator") can freely divide a given amount of resources between himself and a recipient who has no choice but to accept the offer¹⁰, which is why DG is often employed to measure altruistic behaviour. A mean giving rate of about 30% reported in the DG literature⁸ indicates that average subjects show a generous sharing behaviour at their own expenses. A variation of this paradigm, namely the UG, is constructed in the same manner with the exception that the recipient can deny the proposed share, which results in a loss of gains for both the recipient and the proposer⁹. The UG is therefore thought to reflect the individual's ability to understand fairness norms in social partners and to adjust the giving behaviour accordingly. This is supported by studies showing that recipients decline offers lower than 30% and that proposers offer a larger proportion (40%) of the pie to the responder in a UG than in the DG¹¹. The third paradigm, PD, implements a situation where the participant has the opportunity to either cooperate or defect in a scenario involving another player^{7,12}. In a typical paradigm, participants are presented with a scenario in which cooperation leads to greatest collective benefit, while the individual advantage

¹Department of Clinical Psychology and Psychotherapy, University of Tübingen, Tübingen, Germany. ²LEAD Graduate School & Research Network, University of Tübingen, Tübingen, Germany. Correspondence and requests for materials should be addressed to J.S. (email: jonathan.scheeff@uni-tuebingen.de)

is higher if one person defects and the other does not. In the latter situation, the player who did not defect faces substantial disadvantages. This dilemma between self-interest and collective outcome is assumed to measure cooperation between individuals, with average cooperation rates around 50%¹³ (note that cooperation varies in this paradigm, depending on the used payoff matrix).

There is a large body of research on different aspects of prosocial behaviour in healthy individuals using some variation of these economic games. Recently, these methods have also gained growing attention in clinical psychology where a considerable proportion of disorders are characterized by interactional impairments that are usually difficult to quantify^{14,15}. This line of research provides increasing evidence that antisociality and psychopathic traits in particular may be associated with specific deficits in prosocial behaviour. The construct of psychopathy has been introduced to describe a subpopulation within the antisocial spectrum which is characterized by specific traits in the interpersonal (e.g., manipulateness), affective/cognitive (e.g. callous-unemotional traits), and behavioural domain (antisocial behaviour)¹⁶. Previous research which employed the DG showed associations between psychopathic traits and attenuated sharing behaviour in both community and incarcerated samples^{17–19}. However, decreased proposals in the UG have not been related to psychopathic traits^{18,20}, suggesting more sophisticated behaviour when unfair shares can be punished compared to the DG. Studies that investigated cooperative behaviour by employing the PD reported less cooperation in community samples with elevated antisocial behaviour and psychopathic traits^{17,20–23}, while the findings for incarcerated samples are less consistent^{24,25}. In summary, previous research suggests that while psychopathic and antisocial individuals show attenuated altruism, the responsiveness to descriptive norms appears to be intact, which indicates an understanding of social norms and an ability to adapt their behaviour in order to maximize personal gains. However, cooperation remains an understudied phenomenon in these populations and the psychological mechanisms that govern the decision to cooperate are poorly understood.

The frequency of interactions and the identity of social partners are two important determinants for everyday cooperation. These characteristics of social interactions, namely the affiliation of interaction partners and the number of expected interactions, represent established psychological factors that have been shown to motivate cooperative behaviour^{26,27}. In-group members have been known to be favoured over out-group members²⁸ and cooperation increases for repeated as compared to single interactions due to potential reciprocity⁷. Further, individuals show an even stronger in-group favouritism when their own outcomes depend on their partner's behaviour²⁷. First attempts to bridge these psychological concepts with antisocial characteristics demonstrate an enhanced in-group favouritism for students with elevated psychopathic traits^{29,30}. Further, the absence of cues about future interactions predicted non-cooperation in a PD for students scoring high on psychopathy²³. However, similar studies in clinical or incarcerated samples are pending.

To investigate cooperative behaviour in antisocial individuals and to test the role of group affiliation and the number of expected interactions in motivating cooperation, we used the Give Some Dilemma (GSD) paradigm^{31,32} in a sample of incarcerated violent offenders and control participants. In this task, participants had an amount of 10 monetary units (MUs) which they could divide between themselves and a counterpart. Group affiliation between the participant and the counterpart (in-group vs. out-group) was established using the minimal group paradigm prior to the experiment³³. To test for the effects of the number of expected interactions, we manipulated the number interactions with the same counterpart (single-trial interactions vs. repeated-trial interactions with the same counterpart) during GSD, resulting in 60 trials in total (20 single-trial interactions and a total of 40 trials in the repeated interaction condition). Based on the previously reported literature we predicted that:

- 1) Violent offenders would show generally reduced cooperation, i.e., share less MUs in GSD, when compared to control participants.
- 2) Cooperation would be modulated by the group affiliation (increased cooperation with in-group members compared to out-group members) and the number expected interactions with the same person (higher cooperation rates in repeated-trial compared to single-trial interactions) in both groups, but to a larger extent in violent offenders.
- 3) Higher psychopathic traits would be associated with lower cooperation rates, in particular for single-trial interactions and for interactions with out-group members.

Results

Participant characteristics. Six control participants were excluded from data analysis due to fully or partially meeting criteria for antisocial personality disorder, resulting in a final sample of 52 violent offenders and 46 controls. In the violent offender group, 35 individuals (67.3%) fulfilled the antisocial personality disorder diagnosis. Twelve violent offenders fulfilled criteria for comorbid substance dependence, one fulfilled criteria for comorbid obsessive-compulsive disorder and seven fulfilled criteria for lifetime major depression episode. Violent offenders were convicted for violent crimes such as aggravated battery, first degree murder, kidnapping, robbery, assault, or threat.

Table 1 displays demographic and clinical sample description for both violent offender and control group. The groups did not differ in terms of years of education and intelligence (*Wiener Matrizen Test 2*, WMT)^{34,35}, but violent offenders were slightly older than control participants. Compared to the controls, violent offenders showed higher Buss-Perry Aggression Questionnaire (BPAQ)³⁶ scores for all subscales and the total score, except for the 'hostility subscale'. Elevated scores in violent offenders for the Hare Self-Report Psychopathy Scale III (SRPS)³⁷ were observed for the subscale 'criminal tendencies' and the total score.

	VO (N = 52)	CTL (N = 46)	Statistics
Demographics			
Age	38.17 (10.04)	33.15 (10.53)	$t(96) = 2.41, p = 0.018$
Education (years)	9.63 (1.28)	9.85 (0.63)	$t(96) = -1.06, p = 0.292$
WMT score	7.35 (3.87)	8.76 (3.77)	$t(96) = -1.83, p = 0.071$
BPAQ			
Physical aggression	22.52 (7.97)	18.80 (6.28)	$t(96) = 2.54, p = 0.013$
Verbal aggression	15.94 (3.80)	14.57 (2.86)	$t(96) = 2.00, p = 0.048$
Anger	16.69 (5.64)	13.87 (3.98)	$t(96) = 2.83, p = 0.006$
Hostility	23.13 (6.56)	21.26 (5.46)	$t(96) = 1.52, p = 0.131$
Total score	78.29 (19.87)	68.50 (14.32)	$t(96) = 2.77, p = 0.007$
SRPS			
Interpersonal manipulation	2.61 (0.53)	2.51 (0.38)	$t(96) = 1.02, p = 0.308$
Callous affect	2.45 (0.58)	2.33 (0.40)	$t(96) = 1.13, p = 0.260$
Erratic lifestyle	3.01 (0.60)	2.84 (0.59)	$t(96) = 1.38, p = 0.170$
Criminal tendencies	2.89 (0.73)	1.80 (0.51)	$t(96) = 8.47, p < 0.001$
Total score	2.74 (0.52)	2.37 (0.35)	$t(96) = 4.06, p < 0.001$

Table 1. Demographic and clinical sample description. *Note.* VO, violent offender group; CTL, control group; WMT, Wiener Matritzen Test 2; BPAQ, Buss-Perry Aggression Questionnaire; SRPS, Hare Self-Report Psychopathy Scale. The data presented in the table refers to means and standard deviations for each measure (in parentheses).

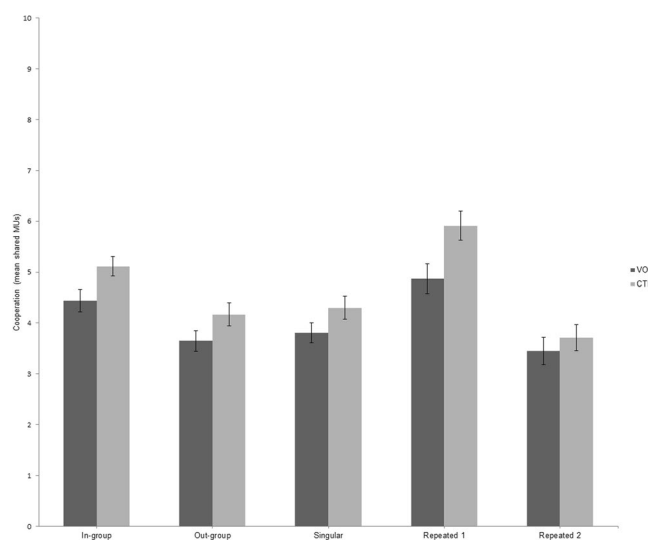


Figure 1. Cooperative behaviour in the Give Some Dilemma as function of shared monetary units (MUs) in violent offenders (VO, $n = 52$) and healthy controls (CTL, $n = 46$); dependent on group affiliation of the other player and number of expected interactions. Singular represents the single-trial interaction condition, Repeated 1 represents the first interaction in the repeated-trial interaction and Repeated 2 represents the second interaction in this block. Error bars indicate standard error from mean.

Cooperative Behaviour. We computed a 2 (group affiliation: in- vs. out-group) \times 3 (number of expected interactions: single-trial interaction vs. first repeated-trial interaction vs. second repeated-trial interaction) repeated measures ANOVA with experimental group (violent offender vs. control) as between-subject variable and shared MUs in the GSD as dependent measure. There was a significant main effect of group ($F(1,96) = 4.620, p = 0.034, \eta_p^2 = 0.046$), indicating lower shared MUs in violent offenders ($M = 4.34, SD = 1.46$) compared to healthy controls ($M = 5.11, SD = 1.39$). Further, we found a significant main effect of group affiliation ($F(1,96) = 60.03, p < 0.001, \eta_p^2 = 0.385$) which revealed higher sharing behaviour with in-group ($M = 5.16, SD = 1.59$) vs. out-group members ($M = 4.25, SD = 1.57$) and a main effect for number of expected interactions ($F(1,96) = 41.29, p < 0.001, \eta_p^2 = 0.301$) indicating more shared MUs in the first repeated-trial interactions ($M = 5.36, SD = 2.09$) compared to single-trial ($M = 4.04, SD = 1.51$) and the second repeated-trial interactions ($M = 3.57, SD = 1.85$; see Fig. 1). No other interaction effects reached significance ($ps > 0.1$). These findings show that the experimental manipulations of group affiliation and expected number of interactions were successfully

	Single	Repeated	In-group	Out-group	SRPS-IM	SRPS-CA	SRPS-EL	SRPS-CT	SRPS-T
Single	—	0.34*	0.66***	0.74***	-0.23	-0.06	-0.27(*)	-0.30*	-0.26(*)
Repeated	0.23	—	0.87***	0.80***	0.05	-0.02	-0.14	-0.05	-0.05
In-group	0.60***	0.79***	—	0.78***	-0.10	-0.07	-0.22	-0.14	-0.15
Out-group	0.71***	0.73***	0.64***	—	0.04	-0.02	-0.23	-0.21	-0.16
SRPS-IM	-0.08	0.18	0.22	-0.06	—	0.69***	0.67***	0.62***	0.86***
SRPS-CA	-0.05	-0.16	0.02	-0.26(*)	0.51***	—	0.55***	0.54***	0.80***
SRPS-EL	0.12	0.22	0.23	0.17	0.32*	0.37*	—	0.77***	0.88***
SRPS-CT	-0.02	-0.13	-0.04	-0.13	0.23	0.29(*)	0.54***	—	0.88***
SRPS-T	0.01	0.05	0.15	-0.07	0.64***	0.69***	0.82***	0.75***	—

Table 2. Correlations between sharing behaviour and diagnostic measures for violent offenders (N = 52) and controls (N = 46). *Note.* Single, shared monetary units (MUs) in single-trial interactions; Repeated, shared MUs in repeated-trial interactions, In-group, shared MUs with in-group counterparts, Out-group, shared MUs with out-group counterparts, SRP-IM, Hare Self-Report Psychopathy Scale (SRP) interpersonal manipulation subscale; SRPS-CA; SRPS callous affect subscale; SRPS-EL, SRPS erratic lifestyle subscale; SRPS-CT, SRPS criminal tendencies subscale; SRPS-T, SRPS total score. The data represented in the table refers to bivariate correlations between the indicated measures for the violent offenders (top) and controls (bottom). *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, (*) $p < 0.10$.

implemented and what that while the violent offenders shared overall less MUs, both groups showed elevated levels of cooperation in the first interaction of the repeated setting and with in-group members.

Influence of psychopathic traits. We found a significant negative correlation between the SRPS subscale ‘criminal tendencies’ as well as a trend toward negative correlations between the SRPS scales ‘erratic lifestyle’ and the total score with cooperation in single-trial interactions for the violent offender group (Table 2). Further, there was trend for a correlation between the SRPS subscale ‘callous affect’ and cooperation with out-group members for the control group. However, these associations were no longer significant after Bonferroni correction.

Post-hoc robustness check: Development of cooperation behaviour. We performed additional post-hoc analyses to describe the development of cooperation over the course of the repeated-trial interaction block (see Supplementary Figs S1 and S2). There is literature showing that individuals change their cooperation behaviour as they receive feedback on their choices³⁸ which is why we performed these post-hoc analyses to see if this was the case in the repeated-trial interaction condition in the GSD. Regression analyses were computed for the cooperation rates in the first interaction for each group, showing an increase of cooperation over time for controls ($b = 0.046$, $t(96) = 6.10$, $p < 0.001$, $R^2 = 0.038$, $F(1, 918) = 37.14$, $p < 0.001$) as well as for the violent offenders ($b = 0.033$, $t(96) = 4.75$, $p < 0.001$, $R^2 = 0.020$, $F(1, 1038) = 22.54$, $p < 0.001$), see Supplementary Fig. S1. This shows that both groups exhibited a slight increase for the first repeated-trial interaction. For the second repeated-trial interaction, the regression analyses showed no change in cooperative behaviour over time, neither for controls ($b = -0.005$, $t(96) = -0.72$, $p = 0.475$, $R^2 = -0.001$, $F(1, 918) = 0.5105$, $p = 0.475$) nor for the violent offenders ($b = 0.005$, $t(96) = 0.69$, $p = 0.492$, $R^2 = -0.001$, $F(1, 1038) = 0.4724$, $p = 0.492$), see Supplementary Fig. S2. Thus, cooperation was stable for the second interaction in the repeated-interaction condition for both groups, despite a lack of a subsequent interaction.

Discussion

The present study aimed to investigate for the first time cooperative behaviour and its’ social-contextual determinants in a population of incarcerated violent offenders using the GSD. We were interested whether violent offenders’ sharing behaviour could be influenced by established social-contextual determinants of cooperative behaviour, i.e., the number of expected interactions and group affiliation. Further, we aimed to investigate psychopathy as a potential moderator for these determinants. The results can be summarized as follows: 1) In line with Hypothesis 1, as compared to controls, the violent offenders shared less MUs with their counterparts, indicating an overall reduced cooperation. 2) In partial support of Hypothesis 2, violent offenders as well as controls showed increased cooperation rates towards in-group members and in repeated-trial interactions. However, these relationships were equivalent in both groups and therefore did not support the assumption of a stronger response to the determinants in violent offenders. 3) In accordance with Hypothesis 3, higher psychopathic traits were associated with lower cooperation in single-trial interactions in the violent offender group. Contrary to this hypothesis, there was no relationship between psychopathic traits and group affiliation in either group, or the number of expected interactions in the control group. In sum, our results show that despite the generally reduced willingness to cooperate, the violent offenders were able to adapt their sharing behaviour to the task demands in the same way as control participants. Importantly, cooperation in violent offenders was determined by the number of expected interactions as well as group affiliation to the same extent as in healthy controls, thus providing evidence for the existence of equivalent psychological mediators. Further, our findings indicate that psychopathic traits in violent offenders are associated with both, more strategic and profit-maximizing behaviour.

Only two studies to date investigated cooperation in incarcerated individuals as compared to healthy controls using the PD paradigm. While one study reported no differences between inmates and control participants²⁵, the other demonstrated lower cooperation rates for incarcerated individuals²⁴. Together with the results of the present

study, the inconsistent findings across the studies warrant some comment. Similar to the study of Mokros *et al.*²⁴, we examined male incarcerated individuals by using a repetitive paradigm, while Khadjavi and Lange²⁵ investigated female prisoners with a one-shot and two-trial paradigm. Therefore, inconsistent findings could be attributed to task characteristics (single vs. repeated interaction) or the gender of the participants, as women have been reported to cooperate more in initial trials than men, but this difference disappears as the interaction continues³⁹. Further research addressing gender specific differences in incarcerated populations is needed. Nevertheless, the current findings add to an existing body of literature documenting lower rates of sharing behaviour in clinical groups as well as community samples with antisocial and psychopathic traits using other economic games^{17,18,20–23}.

To our knowledge, this study is the first to investigate social-contextual determinants of cooperation, i.e., the number of expected interactions and group affiliation, in incarcerated offenders. With regard to the expected number of interactions, our results show that the anticipated frequency of interactions enhances cooperation in both violent offenders and controls, irrespective of the co-player's affiliation with the in- or the out-group. This replicates the findings previously reported for healthy participants⁴⁰ and extends these results to a clinical population. Our experimental design featured aspects from both, the DG (absence of consequences for sharing behaviour as in the single-trial interaction condition) and the UG (possibility for punishment as in the repeated interaction condition) where analogous results have been reported, in particular lower sharing behaviour in the DG^{17–19} but not in the UG^{18,20}. One potential explanation for the higher cooperation rates in the repeated interactions could also be that individuals learn to anticipate the tit-for-tat sharing behavior of their interaction partners over the course of the game. Indeed, cooperation rates increased slightly for the first repeated interaction across experimental trials. Possible motivators for this learning could be building up trust in the reciprocity as well as a more and more strategic behaviour for improving the personal payoff³⁸. In contrast to the latter assumption, cooperation rates remain stable for the second repeated interaction. Further, group affiliation modulates cooperation independent of singular or repeated-trials, which contradicts a mere profit maximizing explanation for the findings behaviour in the presented study. Future studies need to disentangle what exactly individuals were learning in the GSD and which mechanisms enhance cooperation in repeated interactions.

Interestingly, higher psychopathic traits were uniquely associated with lower cooperation in single-trial interactions in the violent offender, but not the control group. This is not consistent with the finding that the absence of cues indicating a possible future interaction predicted non-cooperation in a PD for students scoring high on psychopathic traits²³. Possible explanations for these diverging results in non-incarcerated individuals may be due to differences in implementation (meet someone in real-life versus repeated interactions in a game) and the use of different psychopathy measures. Further, it is not clear whether psychopathic traits can be used in the same way within non-incarcerated samples as with incarcerated populations and therefore, the previous findings should be interpreted with caution. However, in line with this and other previous findings, the relationship between higher psychopathic traits and reduced cooperation in violent offenders indicates a strategic profit-maximization in single-trial interactions with absence of negative consequences. Future studies are needed to disentangle which specific combination of behavioural, antisocial, and psychopathic characteristics account for these behavioural strategies.

Group affiliation is another important determinant of cooperation and our study is the first to provide empirical evidence for a successful application of a minimal group paradigm in incarcerated individuals. The present data supports previous findings demonstrating an in-group bias for healthy individuals²⁷. More importantly, the violent offender group also demonstrated in-group favouritism, revealing no deficits in this key determinant of cooperation. However, contrary to previous studies, psychopathy did not affect the magnitude of the in-group bias^{29,30}. Reasons for this discrepancy may be rooted in differences in study samples (students versus incarcerated offenders) as well as different methods to induce group affiliation (in-group and strangers vs. in-group and out-group). Taken together, the fact that violent offenders enhance their cooperative behaviour in response to group affiliation seems to be encouraging. Nevertheless, due to the unspecific manner in which group affiliation was implemented, we have no knowledge which characteristics of the interaction partner may be driving the identification with group members and whether these characteristics deviate in healthy controls and violent offenders. Thus, despite this first evidence indicating similarities between violent offenders and controls in their cooperative behaviour toward in-group members, it remains to be determined which characteristics evoke group affiliation and how this could be used in a therapeutic context.

The current study has several limitations worth noting. First, the current study has been conducted with an exclusively male population; thus, it remains to be determined whether the effects reported in the current study generalize to female and younger externalizing (e.g. children with conduct disorder) populations. Second, the current study is a laboratory investigation that assessed incarcerated inmates. Therefore, it is not entirely clear whether and to what extent the cooperative behaviour assessed in the current study corresponds with real-life behaviour. Prospective study designs are necessary in order to understand how prosocial acts captured in a laboratory setting relate to real-life behaviour. Lastly, we employed a set number of repeated interactions, which may limit the ecological validity. Future studies should introduce more variation in the number of consecutive interactions in order to more closely mimic real-life scenarios.

The present study is one of the first attempts to address the considerable knowledge gap regarding prosocial behaviour in antisocial individuals. The preliminary results indicate generally lower cooperation in violent offenders but intact responding to recipient and situational characteristics that promote cooperation. Despite pessimism regarding psychological interventions for antisocial individuals⁵ and failure to influence cooperation through emotional feedback in individuals from the antisocial spectrum^{21,41}, the current findings have potential implications for practice, indicating that these factors can be utilized to promote prosocial behaviour. Future research on modulating psychological factors is needed in antisocial individuals in order to understand the nature of the massive behavioural problems in these populations and to develop treatment strategies to alleviate these symptoms.

Method

Participants. Fifty-two incarcerated male violent offenders from cooperating German correctional facilities (Justizvollzugsanstalten Heimsheim, Rottenburg, Hohenasperg) participated in the study. Potential participants were recruited through advertisement via pamphlets and black boards within the facilities. Inclusion criteria were: 18–65 years, primary conviction for violent crimes and sufficient knowledge of the German language. The facility's psychological service contacted interested individuals and scheduled the assessments. Clinical and experimental assessments were carried out in designated rooms of the facility by trained psychologists from our research group. Exclusion criteria for the violent offender group were: history of psychotic-spectrum or bipolar disorders (as assessed by clinical interview), primary conviction for drug-related crime. Fifty-two male participants were recruited via advertisements in newspapers and university's mailing list and served as the control group. Inclusion criteria were: 18–65 years, no self-reported convictions or arrests, no symptoms or full diagnosis of antisocial personality disorder, no history of bipolar or psychosis-spectrum disorder. We matched both groups in terms of years of education through the recruiting process. The study was approved by the Clinical Ethics Committee at the University Hospital Tübingen and was conducted in accordance with the Helsinki Declaration.

Diagnostic and control measures. Self-reported aggression was assessed with a German version of the Buss-Perry Aggression Questionnaire (BPAQ), an instrument that measures trait aggressiveness with 29 items and the subscales physical aggression, verbal aggression, anger, and hostility. The questionnaire contains statements which participants rate on a 5-point Likert scale ranging from 0 (not at all characteristic for me) to 4 (extremely characteristic of me).

Further, we measured psychopathic traits using the Hare Self-Report Psychopathy Scale III (SRPS). The questionnaire contains 64 items and can be divided into four subscales measuring interpersonal manipulation, callous affect, erratic lifestyle, and criminal tendencies. The items contain statements which are rated on a 5-point Likert scale (1 = disagree strongly to 5 = agree strongly).

To control for IQ related cognitive abilities, the short-version of the *Wiener Matrizen Test 2* (WMT) was assessed. The WMT is a non-verbal 18-item test derived from Raven's Progressive Matrices Test that measures problem solving and deductive reasoning. Participants have to match different patterns of matrices to an analogous missing matrix piece by selecting the corresponding part out of eight options.

Current and life-time psychopathology was evaluated with the *Mini International Neuropsychiatric Interview* M.I.N.I.^{42,43} which was administered by trained members of our research group with extensive experience conducting clinical interviews. The M.I.N.I. assesses DSM-IV and ICD-10 criteria for all Axis-I disorders as well as the antisocial personality disorder.

Experimental task. Cooperation was assessed with a modified version of the sequential GSD paradigm^{31,32,40}. We manipulated the number of expected interactions (single-trial interaction vs. repeated-trial interaction) as well as the group affiliation (interaction with in-group vs. out-group member). The dependent variable, cooperative behaviour, was calculated as mean shared MUs for every participant in every condition.

Group affiliation. To manipulate group affiliation, we used the well-established minimal group paradigm³³ by instructing participants before the task that all players including the participants themselves were divided in two groups based on personality characteristics previously assessed in questionnaire measures. The participants were further informed that the group affiliation was indicated by a colour cue (blue or yellow) and that all members of this group were very similar, whereas members of the other group were very different⁴⁴. Subjects were randomly allocated to a group and also to one of the two sets.

Stimuli. 40 photographs depicting males with neutral facial expression were selected from the Radboud Faces Database⁴⁵ and the Karolinska Directed Emotional Faces Database⁴⁶. Two sets of faces were created, where half of the pictures were marked with a blue background and the other half was marked with a yellow background in one set, the other set contained the same pictures with reverse colouring to guarantee that every face served as an in- or out-group member an equal number of times. Colouring was done with GNU Image Manipulation Program (GIMP). Photographs were presented randomly. A stimulus set comprising additional four male faces with neutral expressions was created for practice trials.

Give Some Dilemma task. The participants were informed that they will play a game with a fellow player, who either belonged to the same group (very similar) or to another group (very dissimilar). They received no further information about the identity of the counterpart. Further, they were informed about the opportunity to win MUs, which were later converted into real money (Euro). This additional gain (ranging from 1.80 and 2.10 Euro) to the compensation was implemented to create a more realistic scenario with real-life consequences of cooperative behaviour.

Every trial followed the same pattern: Both players started with ten MUs which they had to divide between themselves and the counterpart at any rate by indicating the number of MUs intended to share via a button press (0–10). Kept MUs counted one fold for the player himself, shared MUs counted twice for the other person. For instance, if a person were to share 4 MUs, the co-player would effectively receive 8 units, while the participant kept the remaining 6. This situation leads to a social dilemma where the individual outcomes are greater to the extent that fewer coins are given away but also both individuals obtain greater outcomes to the extent that they cooperate with each other (give more coins to each other) see³¹.

Participants were told that a random algorithm decided that the fellow player always went first. Each trial started with two profile pictures (see Fig. 2): One on the bottom left which was empty, representing the participant, as well as a neutral photo of a male model on the top right. Both profiles were highlighted either blue or

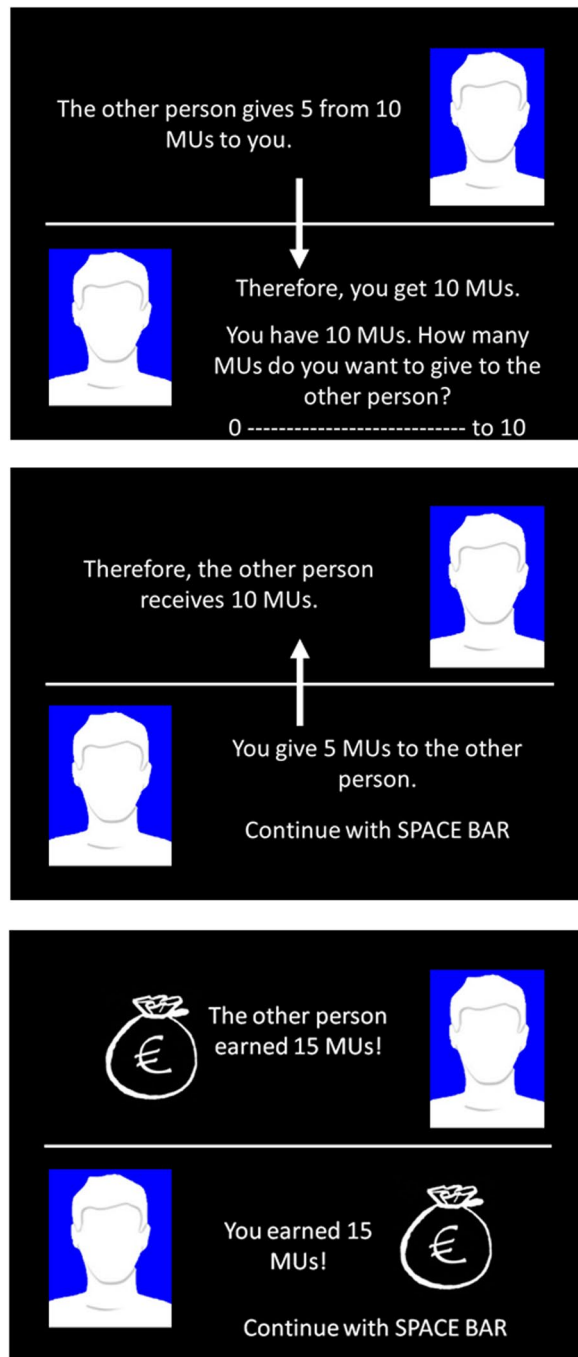


Figure 2. Schematic illustration of a trial in the social dilemma task; MU = monetary unit.

yellow to indicate group membership. The screen was divided by a horizontal bar to simulate two sides of a table. Every trial began with a proposition of the fellow player which amounted 5 MUs in 50% of all trials and 4 or 6 MUs in 25% of the trials. This served to ensure fair offers from the fellow player and to enable the participants to reciprocate less, with an equal number of coins, or more coins⁴⁰. Participants received immediate feedback about their gains and subsequently made their offer which also resulted in an instant feedback of total profits for both players.

Number of expected interactions. In order to manipulate the number of expected interactions, we designed two conditions of the GSD which were administered in two consecutive blocks with a fixed order. The first block comprised single-trial interactions whereas the second block contained repeated-trial interactions (two interactions). In the latter condition, the first interaction started with the same fair offer as described above, whereas the proposal in the second interaction was analogous to the participants share in the first interaction. This reciprocal cooperation strategy is called tit-for-tat which helps efficiently to establish and retain cooperation⁷.

Participants played 20 trials with 20 different fellow players in the first block and 20 repeated interactions with 20 different fellow players (two trials per fellow player), adding up to a total of 60 trials.

Procedure. At the beginning of each assessment, participants provided a written informed consent. The diagnostic interview, WMT, and self-report questionnaires (BPAQ, SRPS) were administered before the subjects were introduced to the GSD task. The full instructions for the presented GSD can be found in the Supplementary methods. They completed three practice trials before Block 1 and one before Block 2. All assessments for the control group were carried out in individual laboratory rooms from the University of Tübingen. Violent offenders were assessed in designated rooms within the facility, where none of facility members was present in the room during testing. The experiment was run on a 14.1" HP notebook with a viewing distance of approximately 40 cm. The experiment was programmed and run in Presentation (Version 16.5, Neurobehavioral Systems).

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on request.

References

- Marsh, A. A. & Blair, R. J. R. Deficits in facial affect recognition among antisocial populations: a meta-analysis. *Neurosci. Biobehav. Rev.* **32**, 454–465, <https://doi.org/10.1016/j.neubiorev.2007.08.003> (2008).
- Dawel, A., O'Kearney, R., McKone, E. & Palermo, R. Not just fear and sadness: meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neurosci. Biobehav. Rev.* **36**, 2288–2304, <https://doi.org/10.1016/j.neubiorev.2012.08.006> (2012).
- Hoppenbrouwers, S. S., Bulten, B. H. & Brazil, I. A. Parsing fear: A reassessment of the evidence for fear deficits in psychopathy. *Psychol. Bull.* **142**, 573–600, <https://doi.org/10.1037/bul0000040> (2016).
- Scott, S., Knapp, M., Henderson, J. & Maughan, B. Financial cost of social exclusion: follow up study of antisocial children into adulthood. *Br. Med. J.* **323**, 1–5, <https://doi.org/10.1136/bmj.323.7306.191> (2001).
- Gibbon, S. *et al.* Psychological interventions for antisocial personality disorder. *Cochrane Database Syst. Rev.*, CD007668–CD007668, <https://doi.org/10.1002/14651858.CD007668.pub2> (2010).
- Chen, X., Sasaki, T. & Perc, M. Evolution of public cooperation in a monitored society with implicated punishment and within-group enforcement. *Sci. Rep.* **5**, 17050, <https://doi.org/10.1038/srep17050> (2015).
- Axelrod, R. & Hamilton, W. D. The evolution of cooperation. *Science* **211**, 1390–1396 (1981).
- Engel, C. Dictator games: A meta study. *Experimental Economics* **14**, 583–610, <https://doi.org/10.1007/s10683-011-9283-7> (2011).
- Güth, W., Schmittberger, R. & Schwarze, B. An experimental analysis of ultimatum bargaining. *J. Econ. Behav. Organ.* **3**, 367–388, [https://doi.org/10.1016/0167-2681\(82\)90011-7](https://doi.org/10.1016/0167-2681(82)90011-7) (1982).
- Forsythe, R., Horowitz, J. L., Savin, N. E. & Sefton, M. Fairness in simple bargaining experiments. *Games Econ. Behav.* **6**, 347–369, <https://doi.org/10.1006/game.1994.1021> (1994).
- Oosterbeek, H., Sloof, R. & Van De Kuilen, G. Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics* **7**, 171–188, <https://doi.org/10.1023/b:exec.0000026978.14316.74> (2004).
- Rapoport, A. & Chammah, A. M. *Prisoner's dilemma: A study in conflict and cooperation*. Vol. 165 (University of Michigan press, 1965).
- Sally, D. Conversation and cooperation in social dilemmas a meta-analysis of experiments from 1958 to 1992. *Ration. Soc.* **7**, 58–92, <https://doi.org/10.1177/1043463195007001004> (1995).
- King-Casas, B. & Chiu, P. H. Understanding interpersonal function in psychiatric illness through multiplayer economic games. *Biol. Psychiatry* **72**, 119–125, <https://doi.org/10.1016/j.biopsych.2012.03.033> (2012).
- Kishida, K. T., King-Casas, B. & Montague, P. R. Neuroeconomic approaches to mental disorders. *Neuron* **67**, 543–554, <https://doi.org/10.1016/j.neuron.2010.07.021> (2010).
- Coid, J. & Ullrich, S. Antisocial personality disorder is on a continuum with psychopathy. *Compr. Psychiatry* **51**, 426–433, <https://doi.org/10.1016/j.comppsy.2009.09.006> (2010).
- Berg, J. M., Liliensfeld, S. O. & Waldman, I. D. Bargaining with the devil: Using economic decision-making tasks to examine the heterogeneity of psychopathic traits. *Journal of Research in Personality* **47**, 472–482, <https://doi.org/10.1016/j.jrp.2013.04.003> (2013).
- Koenigs, M., Kruepke, M. & Newman, J. P. Economic decision-making in psychopathy: a comparison with ventromedial prefrontal lesion patients. *Neuropsychologia* **48**, 2198–2204, <https://doi.org/10.1016/j.neuropsychologia.2010.04.012> (2010).
- Chmura, T., Engel, C. & Englerth, M. At the mercy of a prisoner three dictator experiments. *Appl. Econ. Lett.* **24**, 774–778, <https://doi.org/10.1080/13504851.2016.1226486> (2017).
- Curry, O., Chesters, M. J. & Viding, E. The psychopath's dilemma: The effects of psychopathic personality traits in one-shot games. *Personal. Individ. Differ.* **50**, 804–809, <https://doi.org/10.1016/j.paid.2010.12.036> (2011).
- Johnston, L., Hawes, D. J. & Straiton, M. Psychopathic Traits and Social Cooperation in the Context of Emotional Feedback. *Psychiatry, Psychology and Law* **21**, 767–778, <https://doi.org/10.1080/13218719.2014.893550> (2014).
- Rilling, J. K. *et al.* Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biol. Psychiatry* **61**, 1260–1271, <https://doi.org/10.1016/j.biopsych.2006.07.021> (2007).
- Gervais, M. M., Kline, M., Ludmer, M., George, R. & Manson, J. H. The strategy of psychopathy: primary psychopathic traits predict defection on low-value relationships. *Proc. R. Soc. Lond. B Biol. Sci.* **280**, 20122773, <https://doi.org/10.1098/rspb.2012.2773> (2013).
- Mokros, A. *et al.* Diminished cooperativeness of psychopaths in a prisoner's dilemma game yields higher rewards. *J. Abnorm. Psychol.* **117**, 406–413, <https://doi.org/10.1037/0021-843X.117.2.406> (2008).
- Khadjavi, M. & Lange, A. Prisoners and their dilemma. *J. Econ. Behav. Organ.* **92**, 163–175, <https://doi.org/10.1016/j.jebo.2013.05.015> (2013).
- Wilson, W. Reciprocation and other techniques for inducing cooperation in the Prisoner's Dilemma game. *J. Confl. Resolut.* **15**, 167–195, <https://doi.org/10.1177/00220277101500205> (1971).
- Balliet, D., Wu, J. & De Dreu, C. K. Ingroup favoritism in cooperation: A meta-analysis. *Psychol. Bull.* **140**, 1556–1581, <https://doi.org/10.1037/a0037737> (2014).
- Brewer, M. B. In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychol. Bull.* **86**, 307–324, <https://doi.org/10.1037/0033-2909.86.2.307> (1979).
- Arbuckle, N. L. & Cunningham, W. A. Understanding everyday psychopathy: shared group identity leads to increased concern for others among undergraduates higher in psychopathy. *Soc. Cogn.* **30**, 564–583, <https://doi.org/10.1521/soco.2012.30.5.564> (2012).
- Gillespie, S. M., Mitchell, I. J., Johnson, I., Dawson, E. & Beech, A. R. Exaggerated intergroup bias in economical decision making games: differential effects of primary and secondary psychopathic traits. *PLoS ONE* **8**, e69565, <https://doi.org/10.1371/journal.pone.0069565> (2013).

31. Van Lange, P. A., Ouwerkerk, J. W. & Tazelaar, M. J. How to overcome the detrimental effects of noise in social interaction: the benefits of generosity. *J. Pers. Soc. Psychol.* **82**, 768–780, <https://doi.org/10.1037/0022-3514.82.5.768> (2002).
32. Van Lange, P. A. & Kuhlman, D. M. Social value orientations and impressions of partner's honesty and intelligence: A test of the might versus morality effect. *J. Pers. Soc. Psychol.* **67**, 126–141, <https://doi.org/10.1037/0022-3514.67.1.126> (1994).
33. Tajfel, H., Billig, M. G., Bundy, R. P. & Flament, C. Social categorization and intergroup behaviour. *Eur. J. Soc. Psychol.* **1**, 149–178, <https://doi.org/10.1002/ejsp.2420010202> (1971).
34. Formann, A. K. & Piswanger, K. *Wiener Matrizen-Test. Manual.* (Beltz Test GmbH, 1979).
35. Formann, A. K., Waldherr, K. & Piswanger, K. *Wiener Matrizen-Test 2 (WMT-2): Ein Rasch-skaliertes sprachfreier Kurztest zur Erfassung der Intelligenz.* (Beltz Test GmbH, 2011).
36. Buss, A. H. & Perry, M. The aggression questionnaire. *J. Pers. Soc. Psychol.* **63**, 452, <https://doi.org/10.1037/0022-3514.63.3.452> (1992).
37. Paulhus, D. L., Neumann, C. S. & Hare, R. D. *Manual for the Hare Self-Report Psychopathy Scale.* (Multi-Health Systems, 2012).
38. Burton-Chellew, M. N., Nax, H. H. & West, S. A. Payoff-based learning explains the decline in cooperation in public goods games. *Proc. R. Soc. Lond. B Biol. Sci.* **282**, 20142678, <https://doi.org/10.1098/rspb.2014.2678> (2015).
39. Ortmann, A. & Tichy, L. K. Gender differences in the laboratory: evidence from prisoner's dilemma games. *J. Econ. Behav. Organ.* **39**, 327–339, [https://doi.org/10.1016/S0167-2681\(99\)00038-4](https://doi.org/10.1016/S0167-2681(99)00038-4) (1999).
40. Van Lange, P. A., Klapwijk, A. & Van Munster, L. M. How the shadow of the future might promote cooperation. *Group Process. Intergr. Relat.* **14**, 857–870, <https://doi.org/10.1177/1368430211402102> (2011).
41. Klapwijk, E. T. *et al.* Fairness decisions in response to emotions: a functional MRI study among criminal justice-involved boys with conduct disorder. *Soc. Cogn. Affect. Neurosci.* **11**, 674–682, <https://doi.org/10.1093/scan/nsv150> (2016).
42. Ackenheil, M., Stotz-Ingenlath, G., Dietz-Bauer, R. & Vossen, A. *M.I.N.I. Mini International Neuropsychiatric Interview, German Version 5.0. 0 DSM IV.* (Psychiatric University Clinic Munich, 1999).
43. Lecrubier, Y. *et al.* The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *Eur. Psychiatry* **12**, 224–231, [https://doi.org/10.1016/S0924-9338\(97\)83296-8](https://doi.org/10.1016/S0924-9338(97)83296-8) (1997).
44. Bernstein, M. J., Young, S. G. & Hugenberg, K. The cross-category effect: Mere social categorization is sufficient to elicit an own-group bias in face recognition. *Psychol. Sci.* **18**, 706–712, <https://doi.org/10.1016/j.jrp.2013.04.003> (2007).
45. Langner, O. *et al.* Presentation and validation of the Radboud Faces Database. *Cogn. Emot.* **24**, 1377–1388, <https://doi.org/10.1080/0269930903485076> (2010).
46. Goeleven, E., De Raedt, R., Leyman, L. & Verschuere, B. The Karolinska directed emotional faces: a validation study. *Cogn. Emot.* **22**, 1094–1118, <https://doi.org/10.1080/0269930701626582> (2008).

Acknowledgements

AJ is funded by the LEAD Graduate School & Research Network [GSC1028], a project of the Excellence Initiative of the German federal and state governments. Aside from financial support, no further contributions were made by the funders. We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Tübingen. The authors would like to thank Nina Anger and Julia Graebe for their support in data collection.

Author Contributions

A.J. and M.S. designed the study. J.S. analysed the data and drafted the manuscript. All authors coordinated data collection, contributed to the interpretation of the data and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-35450-z>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018

Appendix D

**Effects of group affiliation on neural signatures of fairness norm violations in antisocial
violent offenders**

Jonathan Scheeff^{1*}, Karsten Rauss² and Michael Schönenberg¹

¹ *Department of Clinical Psychology and Psychotherapy, University of Tübingen, Germany*

² *Institute of Medical Psychology and Behavioral Neurobiology, University of Tübingen,
Germany*

Author Note

*Correspondence concerning this article should be addressed to: Jonathan Scheeff, University of Tübingen, Department of Clinical Psychology and Psychotherapy, Schleichstraße 4, 72076 Tübingen, Germany, Tel.: +49 7071 29-78294, Email: jonathan.scheeff@uni-tuebingen.de

Acknowledgement

The authors would like to thank Annika Boss, Clara Held, Franziska von Helmolt and Marie Heßlinger for their support in data collection. Moreover, we would like to thank the staff of the correctional facilities Hohenasperg and Heimsheim for their support in recruitment of participants and organization of data collection.

Funding

This research was funded by a grant to MS from the German Research Foundation (Scho 1448/3-1). Aside from financial support, the funders made no further contributions.

Conflict of interest

None of the authors have potential conflicts of interest to be disclosed.

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

Abstract

Objective

While norm-violating behavior in antisocial individuals has been widely studied, little is known about how these people react to unfair behavior directed towards them. Previous research yields inconclusive results with some evidence for rational and strategic behavior in antisocial individuals. Electrophysiological correlates as well as socio-contextual factors such as group affiliation that may inform decision making on fairness considerations have not been investigated in previous studies.

Methods

This study compared fairness considerations of antisocial violent offenders (N = 25) and controls (N = 26) by using the Ultimatum Game where one player proposes a split of resources and the other player has to respond by accepting or declining the offer. Group affiliation of the proposer (in- vs. out-group) and fairness of offers (fair vs. unfair) were manipulated.

Results

We found no difference between groups regarding decision behavior. However, healthy participants showed an electrophysiological response to group affiliation, which was attenuated in the violent offender group.

Conclusions

This data suggests intact understanding of social norms in antisocial violent offenders while electrophysiological response pattern may be linked to impaired emotional reactions to expectancy violations.

Significance

This study adds to a body of literature demonstrating neurological and emotional impairments in this population.

Keywords: Fairness Norm Violations, Antisocial Personality Disorder, Ultimatum Game, Aggressive Behavior, Medial Frontal Negativity, Social Decision-making

1. Introduction

Violations of social norms such as cooperation, reciprocity and fairness are characteristics of severe forms of externalizing pathology such as antisocial personality disorder (ASPD). Examples of dissocial acts and social norm violations committed by people with ASPD include extensive criminal activities like theft-related offenses, fraud, and violence. While there is ample research that investigates the precursors, course, and outcome of these norm-violating behaviors in individuals with ASPD (Chapman, Gillespie, & Mitchell, 2018; Compton, Conway, Stinson, Colliver, & Grant, 2005; Petras et al., 2008; Scott, Knapp, Henderson, & Maughan, 2001), surprisingly little is known about how these people react to aggressive and unfair behavior directed towards them.

The ultimatum game (UG) and the dictator game (DG) are prominent paradigms to experimentally investigate fairness norms and reactions to their violation in the laboratory. During the last decades, these games have gained growing attention in clinical research due to their ability to quantify social decision-making and interpersonal impairments in mental disorders. In both games, one participant (the “proposer”) can divide a given amount of resources between himself/herself and another individual (the “responder”). Whereas the responder has no option but to accept in the DG (Forsythe, Horowitz, Savin, & Sefton, 1994), the responder in the UG can deny or accept the proposed share. An accepted offer in the UG is divided as proposed, whereas a rejected offer results in a loss of gains for both the responder and the proposer (Güth, Schmittberger, & Schwarze, 1982). As consequence, a rational proposer in both games should give as little as possible; and a rational responder in the UG should accept any (non-zero) offer, since a small amount of resources is better than nothing.

However, extensive research shows that proposers typically offer a share of around 30% in the DG (Engel, 2011) and around 50% in the UG, while responders in the UG frequently reject offers lower than 30% of the total resources (Camerer, 2003; Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003; Yamagishi et al., 2009). These findings indicate that humans are not rationally maximizing short-term profits, but that emotional reactions to perceived violations of fairness norms play a key role in the UG even if they come at a personal cost (Sanfey et al., 2003). This argument received empirical support from research showing that other factors than rational evaluations such as group affiliation of players can inform economic decision-making. UG studies came to the result that

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

healthy individuals show an in-group favoritism by accepting more offers from in-group as compared to out-group members, even if individuals were categorized into groups based on minimal criteria that are trivial or arbitrary (Brüne et al., 2012; McLeish & Oxoby, 2011; Wang et al., 2017).

Despite a wealth of knowledge about fairness considerations in healthy individuals, less is known about decision making in people with ASPD. Evidence from the DG suggests that these individuals share fewer resources and therefore behave more profit-oriented (Chmura, Engel, & Englerth, 2017; Mayer, Rauss, Pourtois, Jusyte, & Schönenberg, 2018). Psychopaths, a more severely impaired subgroup of ASPD individuals (Coid & Ullrich, 2010), have also been reported to exhibit such reduced sharing behavior in the DG (Berg, Lilienfeld, & Waldman, 2013; Koenigs, Kruepke, & Newman, 2010). However, data from UG experiments with incarcerated individuals as proposers indicates similar behavior of antisocial and psychopathic individuals when compared to healthy controls (Koenigs et al., 2010; Mayer et al., 2018). A possible explanation for these diverging results might be that antisocial individuals exhibit a strong strategic behavior in the UG, where unfair offers can be punished by others (Berg et al., 2013; Curry, Chesters, & Viding, 2011; Mayer et al., 2018).

To date, it is not clear whether this strategic behavior also holds if antisocial individuals play the UG in the role of the responder. Findings from the few existing studies are mixed, with some reporting no differences between incarcerated individuals and controls (Koenigs et al., 2010; Radke, Brazil, Scheper, Bulten, & De Bruijn, 2013) and others suggesting generally higher acceptance rates indicative of more rational behavior in ASPD individuals (Mayer et al., 2018). Studies employing community samples also reported mixed results for participants high in psychopathic traits (Osumi & Ohira, 2010; Vieira et al., 2013), while Koenigs et al. (2010) found even lower acceptance rates for high psychopathic inmates. Furthermore, little is known about non-rational social-contextual factors such as group affiliation, which are known to significantly influence decision-making in healthy individuals (Balliet, Wu, & De Dreu, 2014; Brewer, 1979). A recent study from our own group and data from a community sample suggest comparable or even stronger in-group biases in individuals who exhibit anti-social behaviors while acting as proposers (Gillespie, Mitchell, Johnson, Dawson, & Beech, 2013; Scheeff, Jusyte, & Schönenberg, 2018). Yet, no study to date has investigated the reactions of individuals with ASPD to group context in the role as UG responder.

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

In the present study, we aimed to close these gaps by experimentally manipulating group affiliation of the proposer in an incarcerated sample of antisocial individuals and matched controls. Subjects were responders in the UG and reacted to unfair and fair offers while interacting with in-group or out-group proposers. Participants' brain activity was measured during the experiment using electroencephalography. In particular, we investigated event-related potentials (ERP) in response to fair and unfair offers to elucidate neural and psychological mechanisms underlying decision making in the UG. We focused on the medial frontal negativity (MFN) which is a negative deflection peaking between 250-350ms at fronto-central recording sites in response to unfavorable compared to favorable outcomes (Polezzi et al., 2008), thereby reflecting an early good versus bad evaluation (Hajcak, Moser, Holroyd, & Simons, 2006). Previous research suggested the MFN to reflect an affective evaluation of negative outcomes (Boksem, Tops, Wester, Meijman, & Lorist, 2006; Gehring & Willoughby, 2002; Masaki, Takeuchi, Gehring, Takasawa, & Yamazaki, 2006), violation of social norms (Boksem & De Cremer, 2010) and expectancies (Pfabigan, Alexopoulos, Bauer, & Sailer, 2011; Potts, Martin, Burton, & Montague, 2006; Sato et al., 2005). Therefore, this ERP is typically observed following unfair compared to fair offers in the UG (Boksem & De Cremer, 2010; Hewig et al., 2011). Additionally, Wang et al. (2017) suggested that healthy individuals have higher fairness expectancies towards in-group proposers, resulting in a more pronounced differentiation between fair and unfair offers for in-group as compared to out-group offers.

The MFN is thought to be generated in the anterior cingulate cortex (Gehring & Willoughby, 2002), a brain structure which is crucially involved in decision making in healthy individuals (Bush, Luu, & Posner, 2000; Eisenberger, Lieberman, & Williams, 2003; Sanfey et al., 2003; Shackman et al., 2011). This region is of specific interest in our incarcerated sample due to structural and functional alterations that were reported in the anterior cingulate cortex of antisocial and psychopathic individuals (Glenn, Raine, & Schug, 2009; Glenn, Yang, Raine, & Colletti, 2010; Kiehl, 2006; Kiehl et al., 2001; Koenigs, 2012; Ly et al., 2012; Rilling et al., 2007; Yang & Raine, 2009). Mayer et al. (2018) were the first to investigate the MFN in antisocial individuals by letting participants play an UG. Their study concluded that antisocial violent offenders show a generally attenuated MFN, which might explain a more rational behavior by antisocial and psychopathic individuals, as the MFN reflects emotional responses to expectancy and norm violations.

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

In sum, the current study aims to explore the influence of group membership on antisocial individual's fairness considerations as recipients in the UG. Based on previous studies we expected overall higher acceptance rates in antisocial individuals as compared to controls indicating more rational behavior. In contrast, we hypothesized a similar in-group bias in both groups. Based on previous findings from our group (Mayer et al., 2018), we also expected altered electrophysiological responses in antisocial as compared to healthy individuals. In particular, we anticipated a stronger differentiation between fair and unfair offers for in-group as compared to out-group offers in healthy individuals (Wang et al., 2017), whereas this effect should be reduced in antisocial individuals (Mayer et al., 2018).

2. Methods

2.1. Participants

Twenty-seven incarcerated male antisocial violent offenders (AVOs) from cooperating German correctional facilities (Justizvollzugsanstalten Heimsheim and Hohenasperg) participated in the study. Potential participants were invited to the study by facility members who were not further involved in the study procedure. Inclusion criteria were: 18-65 years, primary conviction for violent crimes and sufficient knowledge of the German language. Trained psychologists from our research group carried out assessments in designated rooms of the facility. Exclusion criteria for the violent offender group were a history of psychotic-spectrum or bipolar disorders (as assessed by clinical interview) and/or a primary conviction for drug-related crime. AVOs were convicted for violent crimes such as first-degree murder, armed or aggravated robbery, aggravated battery, arson, rape, child rape, and hostage taking. In addition, twenty-seven male participants were recruited via online advertisements to serve as control group (CTLs). Inclusion criteria were 18-65 years, no self-reported convictions or arrests, no symptoms or diagnosis of antisocial personality disorder and no history of bipolar or psychosis-spectrum disorder. We matched both groups in terms of years of education through the recruiting process. Participants gave written informed consent and received monetary compensation. The study was approved by the Clinical Ethics Committee at the University Hospital Tübingen and was conducted in accordance with the Helsinki Declaration.

2.2. Clinical measures

Self-reported aggression was assessed with a German version of the Buss-Perry Aggression Questionnaire (BPAQ) (Buss & Perry, 1992). This instrument measures trait aggressiveness with 29 items and the subscales physical aggression, verbal aggression, anger, and hostility. The questionnaire contains statements which participants rate on a 5-point Likert scale ranging from 0 (not at all characteristic for me) to 4 (extremely characteristic of me).

Psychopathy was assessed using the PCL-R (Hare, 2003), which consists of 20 items and measures the construct as a combination of the four facets interpersonal, affective, lifestyle and antisocial. Based on a semi-structured interview and case records, trained raters score each item as 0 = not applicable, 1 = partially applicable, 2 = applicable. The ratings are summed up to a total score between 0 and 40. Independent experts determined PCL-R ratings for the present study as part of the standard forensic diagnostic procedure.

2.3. Procedure

After completing questionnaires, participants were introduced to the UG task. They received information that they would play a game with individuals who had previously participated in the experiment in the role of the UG proposer. All participants of this study played as recipients and were told that, in order to reduce difficulty, only the following offers were possible: 9:1 (9 monetary units (MUs) for the proposer and 1 MU for the recipient), 3:7 and 5:5. Further, they were informed that earned MUs were translated into real money at the end of the experiment and added to the regular reimbursement.

We used a minimal group paradigm to manipulate group membership (Tajfel, Billig, Bundy, & Flament, 1971). Before the task, participants were instructed that all players including the participants themselves were divided in two groups based on similarity concerning answers in the previous questionnaire (BPAQ). Group affiliation was indicated by a color cue (blue or yellow) and participants received further information that all members of same group were very similar, whereas members of the other group were very different. Participants were randomly allocated to one of the groups.

After three practice trials, participants played 160 UG trials. Offers were made by individuals from the same group (in-group condition, 80 trials) or from the other group (out-group condition, 80

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

trials). In each condition, 40 trials were fair offers (5:5) and 40 trials were unfair offers (1:9 and 3:7, 50% each). The experimental conditions (proposer affiliation: in-group vs. out-group; fairness: fair vs. unfair offers) were presented in randomized order, and a break was included after the first half of trials had been completed. Trial numbers per condition were equated across the two blocks.

Each trial started with a fixation cross presented in the middle of the computer screen for 500ms followed by a picture of either an in-group or out-group member for 3000ms (see Figure 1). Subsequently, an offer appeared for 4000ms in the middle of the screen (e.g. 5 / 10). Afterwards, participants had to choose between accepting or rejecting the offer via a corresponding button press. At the end of each trial, participants received feedback about their and the other players earned MUs for 1500ms, followed by a 1000ms inter-trial interval. At the end of the experiment, participants were informed about their overall earnings, which they received in addition to their reimbursement.

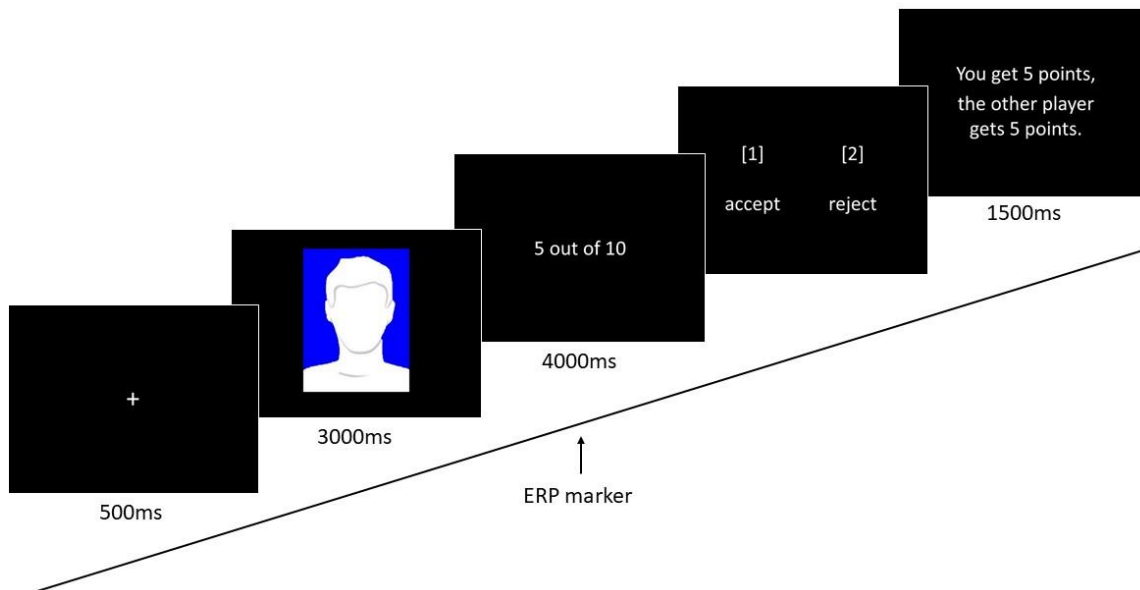


Figure 1. Schematic illustration of an Ultimatum Game trial. Each trial started with a fixation cross followed by a picture of either an in-group or out-group member. Subsequently, a fair or unfair offer appeared, which participants could either accept or reject with a corresponding button press. At the end of a trial, participants received feedback about their and the other players earnings. Event-related potentials (ERPs) were locked to offer presentation.

2.4. Stimuli

Pictures of 80 male models with neutral facial expressions as well as three additional pictures for the practice trials were selected from the Radboud Faces Database (Langner et al., 2010), the Face Database (Gross, 2005) and the Karolinska Directed Emotional Faces (Goeleven, De Raedt, Leyman, & Verschuere, 2008; Lundqvist, Flykt, & Öhman, 1998). In order to adjust the pictures of the different databases, faces were matched for size (495 x 619 pixels) and luminance using Adobe Photoshop CS4® (Adobe Systems Inc., San Jose, USA) and presented in random order against a black background on a 15.4-inch WXGA wide TFT LCD laptop display. For the presentation of visual stimuli, Presentation Software Version 16.4 (Neurobehavioral Systems, USA) was used.

2.5. Electrophysiological recording methods and apparatus

EEG data was recorded from 63 electrodes according to the extended International 10-20 system with a BrainAmp Recorder (BrainProducts, Munich, Germany). Signals were analyzed off-line with BrainVision Analyzer Version 2.1.2.327 software (BrainProducts, Munich, Germany). One additional electrode was placed below the left eye to register vertical eye movements. The recording reference was placed at FCz and the ground electrode was placed at Fpz. Electrode impedances were kept below 10 k Ω . EEG was sampled continuously at a rate of 500 Hz. Continuous EEG data were filtered using a bandpass filter from 0.1 to 30 Hz and re-referenced to the average across all scalp channels. Using Independent Component Analysis, ocular artefact detection and correction were applied to the raw EEG. ERPs were then calculated by extracting epochs of 1000 ms (lasting from -200 ms before to 800 ms after stimulus presentation), baseline-correcting using the 200 ms before stimulus presentation, and averaging epochs per condition.

3. Results

3.1 Participant characteristics

Two participants from the violent offender group and one from the control group were excluded from analysis due to excessive EEG artifacts resulting in a final sample of 25 violent offenders and 26 controls. A demographic and clinical sample description for both, AVOs and CTLs is displayed in Table 1. The groups did not differ in terms of age and years of education. AVOs showed higher aggressiveness as indicated by BPAQ subscales and the total score, except for the verbal aggression

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

and hostility subscale. 21 PCL-R ratings were available for the AVO group with a mean score of 21.81 (SD 9.54). Four PCL-R ratings were not fully determined as part of the standard forensic diagnostic procedure and therefore not available for the current study.

3.2 Behavioral data

Figure 2 displays acceptance rates for different offers and both experimental groups. A 2 (proposer affiliation: in-group vs. out-group) \times 2 (fairness: fair offer vs. unfair offer) repeated measures ANOVA with experimental group (AVOs vs. CTLs) as between-subject variable revealed higher acceptance rates in in-group interactions (mean \pm SE, 0.67 ± 0.04) relative to out-group interactions (0.50 ± 0.04 ; main effect of proposer affiliation, $F(1,49) = 17.16$, $p < 0.001$, $\eta_p^2 = 0.26$). Moreover, acceptance rates for fair offers (0.86 ± 0.03) were higher than for unfair offers (0.44 ± 0.04 ; main effect of fairness, $F(1,49) = 67.04$, $p < 0.001$, $\eta_p^2 = 0.58$). None of the effects reached significance ($p > .100$).

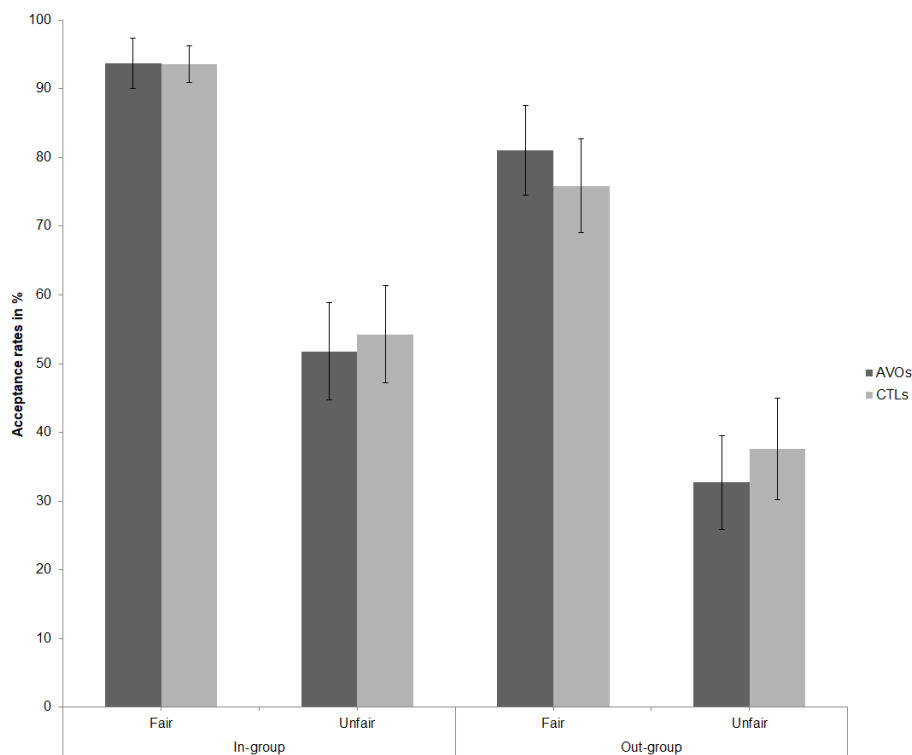


Figure 2. Acceptance rates in the Ultimatum Game for antisocial violent offenders (AVOs, N = 25) and healthy controls (CTLs, N = 26) dependent on group affiliation of the other player (in-group and out-group) and fairness of offers (fair and unfair). Error bars indicate standard error of the mean.

3.3 Electrophysiological data

Figure 3 shows grand average ERP waveforms for AVOs and CTLs. To identify the MFN, we followed the procedures of previous research with similar paradigms (Mayer et al., 2018; Wang et al., 2017) and subtracted grand-average data obtained from CTLs in the fair in-group from those of the unfair in-group condition. The resulting difference waves revealed more negative voltages in the unfair in-group condition at fronto-central electrodes between 230 and 340 ms. Based on visual inspection and the current literature (Fukushima & Hiraki, 2006; Mayer et al., 2018; Wang et al., 2017; Wu, Zhou, van Dijk, Leliveld, & Zhou, 2011), we interpreted this as the MFN and analyzed its amplitude by extracting mean voltage differences (unfair minus fair offers) between 230 and 340 ms at electrode site Cz (Cite Hewig).

A repeated measures ANOVA with proposer affiliation (in-group vs. out-group) as within-subject factor and experimental group (AVOs vs. CTLs) as between-subject variable on MFN amplitudes (unfair minus fair offers) revealed a significant interaction effect between experimental group and proposer affiliation ($F(1,49) = 7.58, p = 0.008, \eta_p^2 = 0.13$). Post-hoc tests showed that MFN amplitudes for in-group offers were more negative (mean \pm SE, $-.29 \pm .23 \mu\text{V}$) as compared to out-group offers (mean \pm SE, $.31 \pm .15 \mu\text{V}$; $t(25) = -2.15, p = 0.041$) in CTLs. This difference was not seen in AVOs (in-group, mean \pm SE, $.22 \pm .27 \mu\text{V}$; out-group offers (mean \pm SE, $-.47 \pm .19 \mu\text{V}$; $t(24) = 1.82, p = 0.081$). Note that there was even a trend for a reversed MFN effect in AVOs, with a numerically more pronounced MFN for in out-group as compared to in-group interactions. In sum, we replicate typical MFN findings for CTLs with a stronger MFN effect in-group offers as compared to out-group offers, whereas this effect was attenuated or possibly reversed in AVOs.

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

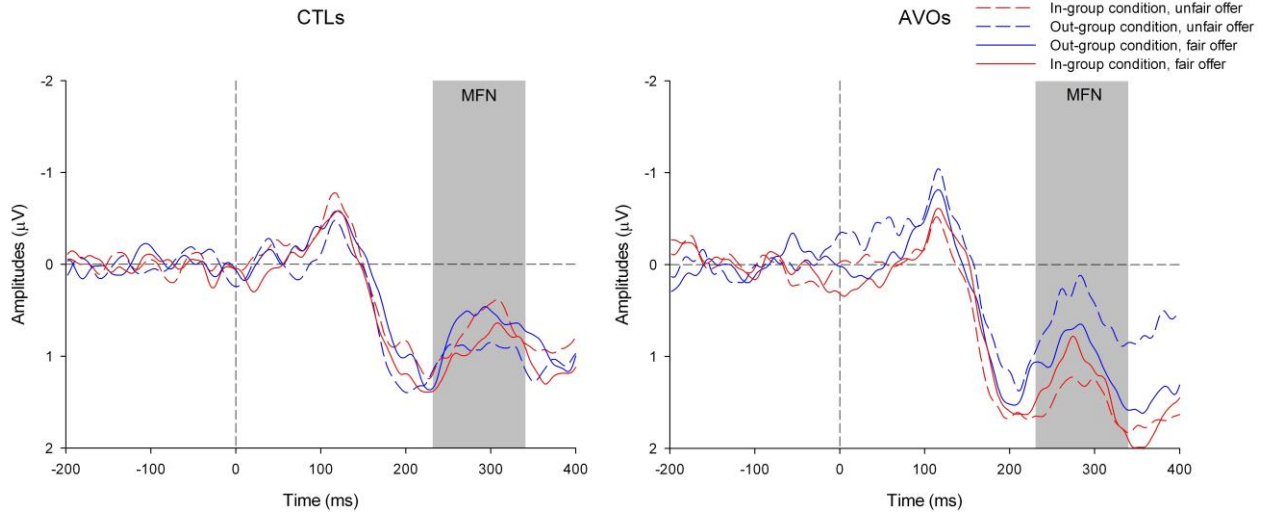


Figure 3. Medial frontal negativity (MFN) amplitudes (230-340ms, grey) at Cz for unfair and fair offers from in- and out-group proposers in the Ultimatum Game as recipient for healthy controls (CTLs, left column) and antisocial violent offenders (AVOs, right column).

3.4 Influence of psychopathic traits in AVOs

We did not find a significant correlation between any PCL-R measure and acceptance rates or MFN responses (Table 2). Note, however, that there was a trend for positive correlations between psychopathic traits and MFN difference waves (where more negative values indicate a stronger effect), indicating an attenuated electrophysiological response with higher psychopathic traits. Further, acceptance rates for fair offers showed a trend towards negatively correlating with psychopathic traits, indicating impaired decision-making with higher psychopathic traits.

4. Discussion

The present study investigated how antisocial violent offenders, as compared to matched controls, react to norm-violating behavior by others towards them. We asked whether their decisions to accept fair and unfair offers were more rational than in healthy individuals, or similarly driven by social norms and socio-contextual factors such as group affiliation. Moreover, we extended this behavioral assessment by measuring electrophysiological responses to fairness norm violations. The results can be summarized as follows: First, acceptance rates of fair and unfair offers from in-group and out-group proposers were similar in AVOs and CTLs. Second, we replicated previous

findings for healthy individuals regarding the MFN response to differences in offer fairness and group affiliation. CTLs showed a stronger MFN effect to in-group offers as compared to out-group offers. Third, and in contrast to the behavioral results, we observed an attenuated MFN response in AVOs. Fourth, psychopathic traits did not significantly influence these results, despite a tendency for stronger attenuation of the MFN response and lower acceptance rates for fair offers with higher psychopathy. In sum, the present study provides evidence for impairments in processing social norm violations on an electrophysiological level in ASPD, while behavioral patterns were similar to the control group.

4.1 Behavioral results

Three previous studies investigated fairness norm considerations in incarcerated populations using the UG paradigm. While two publications reported no difference between inmates and healthy individuals (Koenigs et al., 2010; Radke et al., 2013), the other demonstrated higher acceptance rates indicating strategic, profit-maximizing decision-making (Mayer et al., 2018). The present data support the idea that individuals with ASPD and control subjects behave comparably: both accepted fair offers to a similar extent and showed lower acceptance rates for unfair offers, irrespective of whether proposers were perceived as in-group or out-group members. The experimental manipulation of social context might explain the diverging results of our previous study (Mayer et al., 2018). There, we manipulated social context in terms of whether another human or a computer made the offers, and we found differences between antisocial and healthy individuals. Here, we investigated decision-making in the UG with group affiliation as social context.

Our study provides new evidence that group affiliation modulates fairness norm considerations also in individuals with ASPD. This observation further argues against purely rational behavior of antisocial individuals in social contexts. Rational economic behavior would imply accepting offers irrespective of group membership – a pattern, which we could not identify in the presented data. In contrast, our study confirms previous results demonstrating an in-group bias in antisocial individuals, suggesting no deficits in this key determinant of social decision-making (Scheeff et al., 2018). Another study reported exaggerated in-group preference for students with elevated psychopathic traits (Gillespie et al., 2013). Different populations (students vs. inmates) and a distinct implementation of group affiliation (same or different university vs. similar or dissimilar

individuals) may explain divergences between their and our findings. Taken together, individuals with ASPD seem to understand social rules and adapt their behavior accordingly. Therefore, antisocial behavior in these individuals seems not to be driven by a general misconception of social norms, but by disrespecting known rules occasionally.

Finally, we found a tendency between elevated psychopathic traits and lower acceptance rates towards fair offers, suggesting potential impairments in social decision-making in psychopaths. This finding is in line with previous reports of reduced acceptance rates in incarcerated psychopaths, which were interpreted in terms of a connection between ventromedial prefrontal cortex dysfunction and psychopathy (Koenigs et al., 2010). However, our study found only a tendency for an association and the study of Koenigs et al. (2010) had a small sample size. Future studies should address this issue based on larger samples of individuals scoring high and low in psychopathy.

4.2 Medial frontal negativity

The present findings in healthy controls replicate previous results from an UG study with a group manipulation. Similar to Wang et al. (2017), we show that group affiliation modulates early processes of fairness evaluation as indicated by changes in MFN amplitudes. The MFN is a negative deflection in the ERP following unfair versus fair offers in the UG, thereby reflecting an affective sensibility to violations of fairness expectations. In the current study, the MFN peak was more pronounced when in-group members made offers compared to out-group proposers. Wang et al. (2017) suggested that this MFN modulation by group membership is a result of group-based fairness expectancies. Healthy individuals seem to expect fair offers from in-group members and therefore an unfair offer elicits a more negative MFN amplitude. Yet, CTLs accepted overall more in-group offers irrespective of fairness. This indicates that despite the expectancy violation, they still chose to conform to the social norm of in-group preference. Wang et al. (2017) argued that healthy individuals do so in order to maintain and stabilize in-group integrity when threatened by norm violations of single group members.

In contrast to CTLs, AVOs did not show a differential MFN effect to offers from in- and out-group proposers, indicating attenuated electrophysiological responsiveness. This finding replicates our previous results, which likewise showed an attenuated MFN response in antisocial individuals (Mayer et al., 2018). The present data extend our previous findings in that they highlight the role

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

of socio-contextual factors such as group affiliation. Moreover, we found a tendency for an association between higher psychopathic traits and a stronger attenuation of the MFN response.

We suggest three potential explanations for this phenomenon. First, AVOs have no specific group-based fairness expectancies resulting in an attenuated MFN response. This might be the result of impaired norm understanding, which is in line with the finding that individuals scoring high on psychopathic traits perceive unfair offers as less unfair although acceptance rates were comparable to low psychopathic individuals (Vieira et al., 2013). Yet, our behavioral data argue against this interpretation, as they show a regular rule understanding in ASPD. Second and more likely, individuals with ASPD may understand social norms, but the emotional response to their violation is impaired. In this context, the attenuated MFN response could reflect well-known emotional deficits in antisocial individuals (Dawel, O’Kearney, McKone, & Palermo, 2012; Hoppenbrouwers, Bulten, & Brazil, 2016; Marsh & Blair, 2008) as the MFN mirrors an affective evaluation of outcomes (Boksem et al., 2006; Gehring & Willoughby, 2002; Masaki et al., 2006). Third, it could be that individuals with ASPD have neurophysiological impairments irrespective of norm understanding or specific emotional deficits. It is known that the MFN is generated in the anterior cingulate cortex and this region is known to be altered in antisocial and psychopathic individuals (Glenn et al., 2009; Glenn et al., 2010; Kiehl, 2006; Kiehl et al., 2001; Koenigs, 2012; Ly et al., 2012; Rilling et al., 2007; Yang & Raine, 2009). Therefore, our specific findings might result from (unspecific) structural and functional changes.

Finally, we note that our results show numerical but not statistical evidence for a reversed MFN effect in AVOs. Indeed, AVOs seemed to exhibit a small MFN effect suggesting that they expected fair offers from out-group members. This contrasts with CTLs, who expected fair offers from in-group proposers. This could be due to the fact that group affiliation was manipulated by similarity regarding questionnaire answers in the minimal-group paradigm. Consequently, AVOs might have lower fairness norm expectancies towards individuals who are similar to them. On the other hand, they may expect dissimilar individuals to show fair behavior, as indicated by the more negative MFN amplitude in reaction to unfair offers from those individuals. Notably, if these group-based expectations were indeed present, they did not affect behavioral choices, as AVOs showed clear in-group preferences. Hence, future studies are needed to investigate how perceived group characteristics affect electrophysiological and behavioral responses to fairness norm violations in individuals with ASPD.

4.3 Limitations

The presented study has several limitations that have to be mentioned. First, we solely assessed male individuals with ASPD and therefore future studies should include female participants to investigate whether our findings also hold for ASPD irrespective of gender. Second, while the UG paradigm was developed to mimic social interaction and although the experiment had real-life consequences in terms of additional monetary, its ecological validity is limited. We can only speculate whether the behavior of AVOs in the UG would transfer to real-life scenarios where other factors and stronger incentives might come into play.

4.4 Conclusion

In sum, the present study shows similar understanding of fairness norms in a group context by AVOs as compared to CTLs. This finding is in contrast to the assumption that antisocial individuals exhibit a strategic behavior in economic games. Further, ERP data turned out to be attenuated in AVOs suggesting impairments in early fairness norm evaluation processes. The underlying mechanisms for this diverging behavioral and electrophysiological data needs to be determined by future research.

Author contributions

J.S. and M.S. designed the experiment, J.S. carried out the experiment, J.S., K.R. and M.S. analyzed the data, wrote the paper and reviewed the manuscript.

References

- Balliet, D., Wu, J., & De Dreu, C. K. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, *140*(6), 1556-1581. doi:10.1037/a0037737
- Berg, J. M., Lilienfeld, S. O., & Waldman, I. D. (2013). Bargaining with the devil: Using economic decision-making tasks to examine the heterogeneity of psychopathic traits. *Journal of Research in Personality*, *47*(5), 472-482. doi:10.1016/j.jrp.2013.04.003
- Boksem, M. A. S., & De Cremer, D. (2010). Fairness concerns predict medial frontal negativity amplitude in ultimatum bargaining. *Social Neuroscience*, *5*(1), 118-128. doi:10.1080/17470910903202666
- Boksem, M. A. S., Tops, M., Wester, A. E., Meijman, T. F., & Lorist, M. M. (2006). Error-related ERP components and individual differences in punishment and reward sensitivity. *Brain Res*, *1101*(1), 92-101. doi:10.1016/j.brainres.2006.05.004

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

- Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, 86(2), 307-324. doi:10.1037/0033-2909.86.2.307
- Brüne, M., Tas, C., Wischniewski, J., Welpinghus, A., Heinisch, C., & Newen, A. (2012). Hypnotic ingroup-outgroup suggestion influences economic decision-making in an Ultimatum Game. *Consciousness and Cognition*, 21(2), 939-946. doi:10.1016/j.concog.2012.02.009
- Bush, G., Luu, P., & Posner, M. I. (2000). Cognitive and emotional influences in anterior cingulate cortex. *Trends in Cognitive Sciences*, 4(6), 215-222. doi:10.1016/S1364-6613(00)01483-2
- Buss, A. H., & Perry, M. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63(3), 452. doi:10.1037/0022-3514.63.3.452
- Camerer, C. F. (2003). Strategizing in the Brain. *Science*, 300(5626), 1673-1675. doi:10.1126/science.1086215
- Chapman, H., Gillespie, S. M., & Mitchell, I. J. (2018). Facial affect processing in incarcerated violent males: A systematic review. *Aggression and Violent Behavior*, 38, 123-138. doi:10.1016/j.avb.2017.10.006
- Chmura, T., Engel, C., & Englerth, M. (2017). At the mercy of a prisoner three dictator experiments. *Applied Economics Letters*, 24(11), 774-778. doi:10.1080/13504851.2016.1226486
- Coid, J., & Ullrich, S. (2010). Antisocial personality disorder is on a continuum with psychopathy. *Comprehensive Psychiatry*, 51(4), 426-433. doi:10.1016/j.comppsy.2009.09.006
- Compton, W. M., Conway, K. P., Stinson, F. S., Colliver, J. D., & Grant, B. F. (2005). Prevalence, Correlates, and Comorbidity of DSM-IV Antisocial Personality Syndromes and Alcohol and Specific Drug Use Disorders in the United States: Results From the National Epidemiologic Survey on Alcohol and Related Conditions. *The Journal of Clinical Psychiatry*, 66(6), 677-685. doi:10.4088/JCP.v66n0602
- Curry, O., Chesters, M. J., & Viding, E. (2011). The psychopath's dilemma: The effects of psychopathic personality traits in one-shot games. *Personality and Individual Differences*, 50(6), 804-809. doi:10.1016/j.paid.2010.12.036
- Dawel, A., O'Kearney, R., McKone, E., & Palermo, R. (2012). Not just fear and sadness: Meta-analytic evidence of pervasive emotion recognition deficits for facial and vocal expressions in psychopathy. *Neuroscience & Biobehavioral Reviews*, 36(10), 2288-2304. doi:10.1016/j.neubiorev.2012.08.006
- Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An fMRI study of social exclusion. *Science*, 302(5643), 290-292. doi:10.1126/science.1089134
- Engel, C. (2011). Dictator games: A meta study. *Experimental Economics*, 14(4), 583-610. doi:10.1007/s10683-011-9283-7
- Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in simple bargaining experiments. *Games and Economic Behavior*, 6(3), 347-369. doi:10.1006/game.1994.1021
- Fukushima, H., & Hiraki, K. (2006). Perceiving an opponent's loss: Gender-related differences in the medial-frontal negativity. *Social Cognitive and Affective Neuroscience*, 1(2), 149-157. doi:10.1093/scan/nsl020

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

- Gehring, W. J., & Willoughby, A. R. (2002). The Medial Frontal Cortex and the Rapid Processing of Monetary Gains and Losses. *Science*, 295(5563), 2279-2282. doi:10.1126/science.1066893
- Gillespie, S. M., Mitchell, I. J., Johnson, I., Dawson, E., & Beech, A. R. (2013). Exaggerated intergroup bias in economical decision making games: differential effects of primary and secondary psychopathic traits. *PloS One*, 8(8), e69565. doi:10.1371/journal.pone.0069565
- Glenn, A. L., Raine, A., & Schug, R. A. (2009). The neural correlates of moral decision-making in psychopathy. *Molecular Psychiatry*, 14(1), 5-6. doi:10.1038/mp.2008.104
- Glenn, A. L., Yang, Y., Raine, A., & Colletti, P. (2010). No volumetric differences in the anterior cingulate of psychopathic individuals. *Psychiatry Research: Neuroimaging*, 183(2), 140-143. doi:10.1016/j.psychres.2010.05.009
- Goeleven, E., De Raedt, R., Leyman, L., & Verschuere, B. (2008). The Karolinska directed emotional faces: a validation study. *Cognition and Emotion*, 22(6), 1094-1118. doi:10.1080/02699930701626582
- Gross, R. (2005). Face Databases. In S. Z. Li & A. K. Jain (Eds.), *Handbook of Face Recognition* (pp. 301-327). New York, NY: Springer New York.
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, 3(4), 367-388. doi:10.1016/0167-2681(82)90011-7
- Hajcak, G., Moser, J. S., Holroyd, C. B., & Simons, R. F. (2006). The feedback-related negativity reflects the binary evaluation of good versus bad outcomes. *Biological Psychology*, 71(2), 148-154. doi:10.1016/j.biopsycho.2005.04.001
- Hare, R. D. (2003). *The Hare Psychopathy Checklist Revised (2nd ed.)*. Toronto, Canada: Multi-Health Systems.
- Hewig, J., Kretschmer, N., Trippe, R. H., Hecht, H., Coles, M. G. H., Holroyd, C. B., & Miltner, W. H. R. (2011). Why humans deviate from rational choice. 48(4), 507-514. doi:10.1111/j.1469-8986.2010.01081.x
- Hoppenbrouwers, S. S., Bulten, B. H., & Brazil, I. A. (2016). Parsing fear: A reassessment of the evidence for fear deficits in psychopathy. *Psychological Bulletin*, 142(6), 573-600. doi:10.1037/bul0000040
- Kiehl, K. A. (2006). A cognitive neuroscience perspective on psychopathy: evidence for paralimbic system dysfunction. *Psychiatry Research*, 142(2), 107-128. doi:10.1016/j.psychres.2005.09.013
- Kiehl, K. A., Smith, A. M., Hare, R. D., Mendrek, A., Forster, B. B., Brink, J., & Liddle, P. F. (2001). Limbic abnormalities in affective processing by criminal psychopaths as revealed by functional magnetic resonance imaging. *Biological Psychiatry*, 50(9), 677-684. doi:10.1016/S0006-3223(01)01222-7
- Koenigs, M. (2012). The role of prefrontal cortex in psychopathy. *Reviews in the Neurosciences*, 23(3), 253-262. doi:10.1515/revneuro-2012-0036
- Koenigs, M., Kruepke, M., & Newman, J. P. (2010). Economic decision-making in psychopathy: a comparison with ventromedial prefrontal lesion patients. *Neuropsychologia*, 48(7), 2198-2204. doi:10.1016/j.neuropsychologia.2010.04.012
- Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, 24(8), 1377-1388. doi:10.1080/02699930903485076

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

- Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska directed emotional faces (KDEF)*. Stockholm, Sweden: Department of Clinical Neuroscience, Psychology Section, Karolinska Institutet.
- Ly, M., Motzkin, J. C., Philippi, C. L., Kirk, G. R., Newman, J. P., Kiehl, K. A., & Koenigs, M. (2012). Cortical thinning in psychopathy. *American Journal of Psychiatry*, *169*(7), 743-749. doi:10.1176/appi.ajp.2012.11111627
- Marsh, A. A., & Blair, R. J. R. (2008). Deficits in facial affect recognition among antisocial populations: A meta-analysis. *Neuroscience & Biobehavioral Reviews*, *32*(3), 454-465. doi:10.1016/j.neubiorev.2007.08.003
- Masaki, H., Takeuchi, S., Gehring, W. J., Takasawa, N., & Yamazaki, K. (2006). Affective-motivational influences on feedback-related ERPs in a gambling task. *Brain Res*, *1105*(1), 110-121. doi:10.1016/j.brainres.2006.01.022
- Mayer, S. V., Rauss, K., Pourtois, G., Jusyte, A., & Schönberg, M. (2018). Behavioral and electrophysiological responses to fairness norm violations in antisocial offenders. *European Archives of Psychiatry and Clinical Neuroscience*. doi:10.1007/s00406-018-0878-2
- McLeish, K. N., & Oxoby, R. J. (2011). Social interactions and the salience of social identity. *Journal of Economic Psychology*, *32*(1), 172-178. doi:10.1016/j.joep.2010.11.003
- Osumi, T., & Ohira, H. (2010). The positive side of psychopathy: Emotional detachment in psychopathy and rational decision-making in the ultimatum game. *Personality and Individual Differences*, *49*(5), 451-456. doi:10.1016/j.paid.2010.04.016
- Petras, H., Kellam, S. G., Brown, C. H., Muthén, B. O., Ialongo, N. S., & Poduska, J. M. (2008). Developmental epidemiological courses leading to antisocial personality disorder and violent and criminal behavior: Effects by young adulthood of a universal preventive intervention in first- and second-grade classrooms. *Drug and alcohol dependence*, *95*, S45-S59. doi:10.1016/j.drugalcdep.2007.10.015
- Pfabigan, D. M., Alexopoulos, J., Bauer, H., & Sailer, U. (2011). Manipulation of feedback expectancy and valence induces negative and positive reward prediction error signals manifest in event-related brain potentials. *Psychophysiology*, *48*(5), 656-664. doi:10.1111/j.1469-8986.2010.01136.x
- Polezzi, D., Daum, I., Rubaltelli, E., Lotto, L., Civai, C., Sartori, G., & Rumiati, R. (2008). Mentalizing in economic decision-making. *Behavioural Brain Research*, *190*(2), 218-223. doi:10.1016/j.bbr.2008.03.003
- Potts, G. F., Martin, L. E., Burton, P., & Montague, P. R. (2006). When things are better or worse than expected: the medial frontal cortex and the allocation of processing resources. *J Cogn Neurosci*, *18*(7), 1112-1119. doi:10.1162/jocn.2006.18.7.1112
- Radke, S., Brazil, I. A., Scheper, I., Bulten, B. H., & De Bruijn, E. R. (2013). Unfair offers, unfair offenders? Fairness considerations in incarcerated individuals with and without psychopathy. doi:10.3389/fnhum.2013.00406
- Rilling, J. K., Glenn, A. L., Jairam, M. R., Pagnoni, G., Goldsmith, D. R., Elfenbein, H. A., & Lilienfeld, S. O. (2007). Neural correlates of social cooperation and non-cooperation as a function of psychopathy. *Biological Psychiatry*, *61*(11), 1260-1271. doi:10.1016/j.biopsych.2006.07.021
- Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the Ultimatum Game. *Science*, *300*(5626), 1755-1758. doi:10.1126/science.1082976

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

- Sato, A., Yasuda, A., Ohira, H., Miyawaki, K., Nishikawa, M., Kumano, H., & Kuboki, T. (2005). Effects of value and reward magnitude on feedback negativity and P300. *Neuroreport*, *16*(4), 407-411. doi:10.1097/00001756-200503150-00020
- Scheeff, J., Jusyte, A., & Schönenberg, M. (2018). Investigating social-contextual determinants of cooperation in incarcerated violent offenders. *Scientific Reports*, *8*(1), 17204. doi:10.1038/s41598-018-35450-z
- Scott, S., Knapp, M., Henderson, J., & Maughan, B. (2001). Financial cost of social exclusion: follow up study of antisocial children into adulthood. *British Medical Journal*, *323*(7306), 1-5. doi:10.1136/bmj.323.7306.191
- Shackman, A. J., Salomons, T. V., Slagter, H. A., Fox, A. S., Winter, J. J., & Davidson, R. J. (2011). The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nature Reviews Neuroscience*, *12*(3), 154-167. doi:10.1038/nrn2994
- Tajfel, H., Billig, M. G., Bundy, R. P., & Flament, C. (1971). Social categorization and intergroup behaviour. *European Journal of Social Psychology*, *1*(2), 149-178. doi:10.1002/ejsp.2420010202
- Vieira, J. B., Almeida, P. R., Ferreira-Santos, F., Barbosa, F., Marques-Teixeira, J., & Marsh, A. A. (2013). Distinct neural activation patterns underlie economic decisions in high and low psychopathy scorers. *Social Cognitive and Affective Neuroscience*, nst093. doi:10.1093/scan/nst093
- Wang, Y., Zhang, Z., Bai, L., Lin, C., Osinsky, R., & Hewig, J. (2017). Ingroup/outgroup membership modulates fairness consideration: neural signatures from ERPs and EEG oscillations. *Scientific Reports*, *7*, 39827. doi:10.1038/srep39827
- Wu, Y., Zhou, Y., van Dijk, E., Leliveld, M. C., & Zhou, X. (2011). Social comparison affects brain responses to fairness in asset division: an ERP study with the ultimatum game. *Frontiers in Human Neuroscience*, *5*:131. doi:10.3389/fnhum.2011.00131
- Yamagishi, T., Horita, Y., Takagishi, H., Shinada, M., Tanida, S., & Cook, K. S. (2009). The private rejection of unfair offers and emotional commitment. *Proceedings of the National Academy of Sciences*, *106*(28), 11520-11523. doi:10.1073/pnas.0900636106
- Yang, Y., & Raine, A. (2009). Prefrontal structural and functional brain imaging findings in antisocial, violent, and psychopathic individuals: A meta-analysis. *Psychiatry Research: Neuroimaging*, *174*(2), 81-88. doi:10.1016/j.psychresns.2009.03.012

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

Table 1

Participant characteristics and clinical sample description.

	AVO (<i>N</i> = 25)	CTL (<i>N</i> = 26)	Statistics
Demographics			
Age	43.56 (10.72)	44.12 (10.51)	$t(49) = -0.19, p = .853$
Education (years)	9.72 (1.51)	9.57 (0.50)	$t(49) = 0.46, p = .657$
BPAQ			
Physical aggression	22.08 (8.69)	16.77 (5.29)	$t(49) = 2.65, p = .012$
Verbal aggression	14.80 (2.90)	14.35 (2.26)	$t(49) = 0.62, p = .535$
Anger	14.88 (4.97)	11.50 (3.46)	$t(49) = 2.83, p = .007$
Hostility	19.40 (6.71)	16.69 (4.96)	$t(49) = 1.64, p = .107$
Total score	71.16 (18.55)	59.31 (11.57)	$t(49) = 2.75, p = .009$

Note. AVO, antisocial violent offender group; CTL, control group; BPAQ, Buss-Perry Aggression Questionnaire. The data presented in the table refers to means and standard deviations for each measure (in parentheses).

NEURAL SIGNATURES OF FAIRNESS IN OFFENDERS

Table 2

Correlations between acceptance rates and MFN amplitude with psychopathy scores as measured with Psychopathy Checklist Revised (PCL-R) for antisocial violent offenders (N = 21).

PCL-R	In-group	Out-group	Fair	Unfair	In-group MFN	Out-group MFN
Total score	-0.030	-0.025	-0.317	0.092	0.261	0.296
Affect	-0.033	0.097	-0.165	0.109	0.419	.294
Manipulation	-0.086	0.110	-0.229	0.107	0.318	0.377
Lifestyle	-0.008	-0.094	-0.364	0.080	0.194	0.419
Antisocial	0.105	-0.116	-0.237	0.093	-0.069	0.237

Note. Illustrated are bivariate Pearson correlations between indicated measures. MFN, medial frontal negativity; AVOs, antisocial violent offender group; CTL, control group; In-group fair, In-group unfair, Out-group fair, Out-group unfair, acceptance rates for the respective fairness type and group affiliation of the proposer; In-group MFN, Out-group MFN, MFN difference wave between fair minus unfair offers from the in-group or out-group respectively. * $p < 0.05$.