



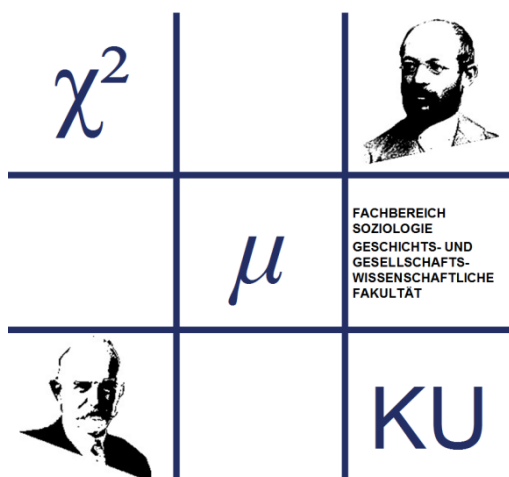
KATHOLISCHE UNIVERSITÄT
EICHSTÄTT-INGOLSTADT

EICHSTÄTTER BEITRÄGE ZUR SOZIOLOGIE

Ein Instrument zur Erfassung allgemeiner Moralvorstellungen (ALLMOR)

Heinz Leitgöb, Stefanie Eifler und Julia Weymeirsch

Nr. 17
Mai 2020



Eichstätter Beiträge zur Soziologie

Die Eichstätter Beiträge zur Soziologie erscheinen in unregelmäßiger Reihenfolge mehrmals im Jahr und können unter der angegebenen Adresse angefordert werden. Für die Inhalte sind allein die jeweiligen Autoren verantwortlich.

Redaktion: Anke Regensburger, Sekretariat des Lehrstuhls für Soziologie und empirische Sozialforschung

Kontakt: Katholische Universität Eichstätt-Ingolstadt

Kapuzinergasse 2
85072 Eichstätt
+49 8421 93 21243
anke.regensburger@ku.de

<http://www.ku.de/ggf/soziologie/eichstaetter-beitraege-zur-soziologie/>

Ein Instrument zur Erfassung allgemeiner Moralvorstellungen (ALLMOR)

Heinz Leitgöb, Stefanie Eifler und Julia Weymeirsch

Korrespondenzanschrift:

Dr. Heinz Leitgöb
Katholische Universität Eichstätt-Ingolstadt
Lehrstuhl für Soziologie und empirische Sozialforschung
Ostenstr. 26
85072 Eichstätt

heinz.leitgoeb@ku.de

Abstract:

In der kriminalsoziologischen bzw. kriminologischen Forschung repräsentieren Moralvorstellungen der handelnden Akteure ein bedeutsames, allerdings auch vernachlässigtes Element der Erklärung kriminellen Handelns. Vielfach werden Moralvorstellungen und Normorientierungen begrifflich gleichgesetzt. Im Rahmen der empirischen Prüfung von Moralitätshypothesen werden vornehmlich *Deviant Morality Scales* zum Einsatz gebracht, die allerdings aus mehreren Gründen dem Anspruch nicht gerecht werden, individuelle Moralvorstellungen valide zu messen. Deshalb verfolgt der Beitrag das Ziel, mit ALLMOR ein theoretisch fundiertes Messinstrument für allgemeine individuelle Moralvorstellungen zu entwickeln. Konkret soll in Anlehnung an die Differentierung in *ethics of being* und *ethics of doing* (Abend 2012) ein zweidimensionales Instrument mit einer seins- (Smor) und einer handlungsorientierten (Hmor) Komponente allgemeiner Moralvorstellungen vorgelegt werden.

Keywords:

Allgemeine Moralvorstellungen, kriminelles Handeln, Skalenkonstruktion, Instrumentenentwicklung, Moral Foundations Theory

1 Einleitung

In der kriminalsoziologischen bzw. kriminologischen Forschung repräsentieren allgemeine Moralvorstellungen der handelnden Akteure ein bedeutsames, allerdings vernachlässigtes Element der Erklärung delinquenten Handelns (z.B. Antonaccio & Tittle, 2008; Hitlin & Vaisey, 2013; Wikström, 2010). Die von Brauer und Tittle (2017, S. 818) als ‚*Morality Theories*‘ bezeichneten Ansätze mit entsprechender Ausrichtung – allen voran die im zentraleuropäischen Forschungsraum vielbeachtete *Situational Action Theory* (SAT; z.B. Wikström et al., 2012)¹ – vertreten im Grundsatz die Ansicht, dass Akteure im Zuge ihrer Entscheidung für (oder gegen) delinquentes Handeln neben Prinzipien der instrumentellen Steuerung (überblicksartig z.B. Eifler & Leitgöb, 2018; Leitgöb & Eifler, 2018; Matsueda, 2013) auch maßgeblich von ihren Moralvorstellungen geleitet werden.²

Jene moralitätstheoretischen Ansätze, die auf dem Konzept des *moralischen Filters* beruhen (z.B. Paternoster & Simpson, 1996; Trasler, 1993; Wikström et al., 2012), gehen zum Teil noch deutlich über die Kodeterminationsannahme von Handlungsentscheidungen durch Kosten-Nutzen-Abwägungen und Moralvorstellungen hinaus. Konzeptionell handelt es sich beim moralischen Filter um einen Wahrnehmungsfiler, der Akteure mit ausgeprägten Moralvorstellungen (basierend auf der Internalisierung und Übernahme hoher moralischer Standards) *automatisch* davon abhält, delinquentes Handeln – auch in Situationen mit kriminalitätsbegünstigenden Gelegenheitsstrukturen (z.B. Cohen & Felson, 1979) – überhaupt als potentielle Handlungsalternative zu perzipieren.³ In diesem Sinne argumentiert Trasler (1993, S. 315): „The ‚moral‘ individual does not admit to his decision set considerations that are inconsistent with his moral stance.“ Diese Akteure stellen Konformität demzufolge im Einklang mit ihren Moralvorstellungen *unbewusst* über die selektive Wahrnehmung der in einer Situation möglichen Handlungsalternativen sicher. Da delinquentes Handeln somit erst gar nicht in das Choice Set (Menge der in einer Situation von einem Akteur perzipierten potentiell ausführbaren Handlungsoptionen) eingehen und einer deliberativen Evaluation auf instrumenteller Grundlage unterzogen werden kann, ist Kriminalität für diese Akteure tatsächlich *undenkbar* (Wikström, 2006; Wikström, 2010).

¹Ferner verweisen Brauer und Tittle (2017) in diesem Zusammenhang auf das *Modell der Frame Selektion* (MFS; einschlägig z.B. Eifler, 2009; Kroneberg et al., 2010; Schulz & Kroneberg, 2018) und die *Institutional Anomie Theory* (IAT; z.B. Messner & Rosenfeld, 2012).

²Die ‚weite‘ Fassung der *Rational Choice Theory* (RCT; z.B. Opp, 1999) erlaubt sogar die theoretische Integration von Moralvorstellungen in das handlungsleitende Kosten-Nutzen Modell. Auf Grundlage der *Subjective Expected Utility Theory* (Savage, 1954) können unter bestimmten Annahmen (siehe dazu z.B. Kroneberg et al., 2010; Opp, 1999, 2013) kognitive Dissonanz (Festinger, 1957) sowie Scham- und Schuldgefühle, die durch das Handeln entgegen der eigenen Moralvorstellungen aufgebaut werden würden, als erwartete ‚psychologische Kosten‘ verstanden werden (z.B. Grasmick & Bursik, 1990; Grasmick & Green, 1980; Nagin & Pogarsky, 2001). Diese werden von den Akteuren im Rahmen der instrumentellen Bewertung perzipierter krimineller Handlungsalternativen mitberücksichtigt.

³Ein vergleichsweise elaboriertes moralisches Filterkonzept legt die SAT vor (Brauer & Tittle, 2017). Es berücksichtigt neben den individuellen Moralvorstellungen der Akteure auch den situativ wahrgenommenen moralischen Kontext, der sich aus den innerhalb des handlungsrelevanten sozialen Settings dominierenden moralischen Standards konstituiert (zu den Details siehe z.B. Wikström, 2010; Wikström et al., 2012). Der moralische Kontext ist allerdings nicht Gegenstand des vorliegenden Beitrags.

Die traditionell erfahrungswissenschaftliche Ausrichtung der einschlägigen Forschungsperspektive wirft nun die Frage auf, wie die angedeutete theoretische Relevanz des Moralitätskonzepts zur Erklärung kriminellen Handelns einer substantiellen empirischen Prüfung unterzogen werden kann. Eine solche setzt die Existenz eines tragfähigen Messinstruments voraus, das das zuvor präzise definierte Phänomen valide in einem festgelegten Zahlenraum abbildet. Bislang liegt allerdings kein Instrument vor, das diesem Anspruch genügt. Vor diesem Hintergrund wird für das interessierende Konstrukt ‚individuelle allgemeine Moralvorstellungen‘ von Akteuren zunächst eine Definition mit anschließender Konzeptspezifikation erarbeitet (Abschnitt 2). Im Anschluss (Abschnitt 3) erfolgt eine kritische Reflexion der *Deviant Morality Scales*⁴ (DMS) als dem gegenwärtig dominierenden Messkonzept für individuelle Moralvorstellungen im Rahmen der empirischen Prüfung entsprechender SAT-Annahmen (Pauwels et al., 2018). Das eigentliche Ziel des Beitrags ist es schließlich, ein auf den theoretischen Ausführungen in Abschnitt 2 basierendes Messinstrument für allgemeine Moralvorstellungen (Akronym: ALLMOR) zu entwickeln, das die in Abschnitt 3 aufgezeigten Limitierungen der DMS überwindet. In Abschnitt 4 wird die methodische Vorgehensweise im Zuge der Skalenkonstruktion, der empirischen Entwicklung und der anschließenden Prüfung der Messeigenschaften des gesamten Instruments sowie dessen Validierung vorgestellt. Die empirischen Ergebnisse werden in Abschnitt 5 aufbereitet. Im finalen Abschnitt 6 erfolgt eine Zusammenfassung und Diskussion der Befunde vor dem Hintergrund der Zielsetzung des Beitrags.

2 Allgemeine Moralvorstellungen – Definition und Konzeptspezifikation

Obwohl die Bedeutung der präzisen theoretischen Konzeptualisierung eines interessierenden (sozialen) Phänomens *vor* der Entwicklung eines entsprechenden Messinstruments offensichtlich ist, wird gerade dieser Aspekt vielfach vernachlässigt (z.B. Billiet, 2016; Saris & Gallhofer, 2007). Die Herausforderung, ein stabiles theoretisches Fundament zugrunde zu legen, offenbart sich im vorliegenden Fall bereits unmittelbar beim Versuch, eine allgemein akzeptierte Nominaldefinition für den Begriff ‚Moralität‘ vorzulegen. So verweisen Gert und Gert (2017) darauf, dass zumindest seit Beginn des 20. Jahrhunderts explizite Versuche, Moralität zu definieren eine Seltenheit darstellen. In ihrem richtungsweisenden Beitrag zur philosophischen Debatte erörtern sie das differenzierte Verständnis des Moralitätsbegriffs, und zwar in einem (i) *deskriptiven* und (ii) *normativen* Sinn (siehe dazu auch bereits Frankena, 1966). Während sich Moralität im deskriptiven Sinn auf “certain codes of conduct put forward by a society or a group (such as a religion), or accepted by an individual for her own behavior” bezieht, wird unter dem normativen Moralitätsverständnis “a code of conduct that, given specified conditions, would be put forward by all rational persons” subsumiert (Gert & Gert, 2017, Online-Ressource, daher keine Seitenreferenz).

(i) Das *deskriptive* Moralitätsverständnis bedarf eines konkreten empirischen Bezugs-

⁴Hierbei handelt es sich um ein Messkonzept, das auf dem Prinzip der expliziten Abfrage basiert, für wie falsch die Befragten die Ausübung unterschiedlicher devianter bis delinquenter Handlungen erachten. Für weitere Ausführungen siehe Abschnitt 3.

rahmens (z.B. einer spezifischen Gesellschaft oder sozialen Gruppe), der den Geltungsbereich dieser moralischen *codes of conduct* (Verhaltensgrundsätze bzw. -normen) begründet.⁵ Verhaltensgrundsätze im Allgemeinen werden von Luco (2014, S. 313) als handlungsregulierende Regeln definiert, die die Ausführung bzw. Unterlassung von Handlungen ge- bzw. verbieten. Sie unterscheiden sich von Gesetzen als dem zentralen staatlichen Instrument der Verhaltensregulierung nicht darin, dass sie weniger konkret oder objektiv, sondern vielmehr weniger klar und monolithisch sind (Cooper, 1966, S. 20). So repräsentierten Gesetze, im Gegensatz zu Verhaltensgrundsätzen, präzise formulierte und kodifizierte Normen, deren Einhaltung von dazu legitimierten Institutionen überwacht und bei Missachtung auf Grundlage eines gleichermaßen festgelegten Strafenkatalogs sanktioniert werden. Weiterhin besitzen Gesetze in ihrem Geltungsbereich universelle Gültigkeit, während in großen sozialen Entitäten wie Gesellschaften nach deskriptivem Moralitätsverständnis mehrere Gruppen mit unterschiedlichen, einander sogar widersprechenden Sets an Verhaltensgrundsätzen existieren können. Daraus ergibt sich unmittelbar die Konsequenz, dass im Rahmen einer deskriptiven Moralitätsdefinition notwendigerweise auch spezifiziert werden muss, *welche* Verhaltensgrundsätze als moralisch klassifiziert werden (Gert & Gert, 2017). Diese werden in der Folge als moralische Verhaltensgrundsätze (MVG) bezeichnet.

(ii) Das *normative* Moralitätsverständnis bezieht sich im Gegensatz dazu auf *universalistische* MVGs, die unter den gleichen Bedingungen kontextunabhängig von jedem *moral agent* akzeptiert werden. Als solcher wird allgemein ein Akteur bezeichnet, der “certain intellectual and volitional conditions, almost always including the condition of being rational” erfüllt (Gert & Gert, 2017, Online-Ressource, daher keine Seitenreferenz).⁶ Nicht alle Akteure sind demzufolge auch *moral agents*. Eine Ausnahme stellen z.B. kleine Kinder dar, da sie kognitiv (noch) nicht in der Lage sind, MVGs semantisch sowie pragmatisch zu erfassen, ihnen zu entsprechen und ihr Handeln daran auszurichten (Haksar, 1998). Ferner argumentieren Gert und Gert (2017), dass auch nicht alle

⁵Insgesamt dominierte die deskriptive Perspektive das wissenschaftliche Moralitätsverständnis des 20. Jahrhunderts (Gert & Gert, 2017). Dies trifft insbesondere auf die Soziologie zu, die zudem, aufbauend auf Marx, Simmel, Durkheim und Weber, eine starke kollektivistische Ausrichtung aufweist (Powell, 2010). Exemplarisch sei Durkheim (1992 [1893]) erwähnt, der alles als moralisch bezeichnet, das eine Quelle von Solidarität repräsentiert und Akteure dazu veranlasst, ihre Handlungen nicht auf der Grundlage egoistischer Motive (individuelle Bedürfnisbefriedigung) auszuwählen. Insgesamt zielt die soziologische Auseinandersetzung, deren Intensität seit den genannten Klassikern stark variiert (z.B. Hitlin & Stephen, 2010), maßgeblich darauf ab, was in unterschiedlichen sozialen und historischen Kontexten als *moralisch klassifiziert* wird bzw. wurde (Hitlin & Vaisey, 2013). Demgegenüber fokussiert die (Sozial-)Psychologie auf “*inside-the-mind stuff*” (Haidt, 2012, S. 314) wie etwa *moral emotions* (Verachtung, Wut, Abscheu, Scham, Schuld und Verlegenheit; siehe z.B. Haidt, 2003; Tangney et al., 2007). Für eine Diskussion der Unterschiede zwischen der soziologischen und der psychologischen Perspektive siehe Hitlin und Vaisey (2013) sowie Whiteley (1960). Ferner sei für die Differenzierung von Moralität als *sozialem* und *individuellem* Phänomen auf Cooper (1966) verwiesen.

⁶Für weitere Ausführungen zum *moral agent* siehe z.B. Haksar (1998) und Mayo (1968). Im Kontext des philosophischen Zugangs zu künstlicher Intelligenz und maschinellem Lernen hat sich in den letzten Jahren auch der Begriff des *artificial moral agents* etabliert. Dieses Themenfeld ist für den vorliegenden Beitrag jedoch nicht von Bedeutung. Interessierte seien dennoch auf die Arbeiten von Allen et al. (2000), Allen et al. (2005) und Johnson (2006) verwiesen.

Verhaltensgrundsätze, die von jedem *moral agent* akzeptiert werden, automatisch moralischer Natur sind. Sie können vielmehr aus alternativen Gründen, z.B. Vernunft oder Rationalität, akzeptiert werden. Weitgehende Einigkeit besteht unter den *moralischen Realisten*⁷ jedoch darin, dass Kant's *Kategorischer Imperativ* (Kant, 1990 [1797], 1961 [1787]) als Idealbeispiel für einen solchen universalistischen MVG gilt.

Zum Zwecke der Formulierung einer tragfähigen Arbeitsdefinition für allgemeine Moralvorstellungen, die der Entwicklung eines entsprechenden Messkonzepts als theoretisches Fundament dienen soll, erscheint es zielführend, mit den MVGs die zentrale Gemeinsamkeit der beiden Perspektiven aufzugreifen. Hinsichtlich der Reichweite ihres Geltungsbereichs muss unter Validitätsgesichtspunkten notwendigerweise *zumindest* die soziale Entität eingeschlossen sein, für die das Instrument konzipiert wird. Dementsprechend muss der Geltungsbereich der MVGs jedenfalls den angestrebten Anwendungsbereich des Instruments abdecken (zu den Details für den vorliegenden Fall siehe die Abschnitte 4.1 und 4.2). Eine solche pragmatische Sichtweise schließt die Existenz universalistischer MVGs selbstverständlich nicht aus, sie setzt sie lediglich nicht voraus.

Auf Grundlage der bisherigen Ausführungen lässt sich schließlich die folgende Definition vorlegen: Allgemeine Moralvorstellungen sind von Akteuren internalisierte und in der Folge auch akzeptierte (für sich übernommene) mentale Repräsentationen genereller moralischer Verhaltensgrundsätze, die deren Handlungssteuerung beeinflussen. Dieser Prozess kann bewusst oder unbewusst (von den Akteuren nicht aktiv wahrgenommen und folglich im automatischen Modus prozessiert) ablaufen. ‚Generell‘ ist in diesem Zusammenhang nicht mit ‚universell‘ gleichzusetzen, sondern im Sinne von ‚abstrakt‘ zu verstehen. Doch gerade ihre generelle Anlage erlaubt es MVGs in den unterschiedlichsten konkreten Situationen maßgeblich zur Entscheidungsfindung zwischen perzipierten Handlungsalternativen beizutragen.

Im Rahmen der Konzeptspezifikation stützen wir uns maßgeblich auf die von Williams (1985) eingeführte und von Abend (2012) prominent vertretene Unterscheidung zwischen *dünnen* (*thin*) und *dicken* (*thick*) Konzepten von Moralität. Während *dünne* Konzepte ein Objekt (z.B. eine Handlung, ein Individuum) lediglich evaluieren, dieses aber nicht beschreiben, beinhalten *dicke* Konzepte beides. D.h. sie überwinden die in der einschlägigen psychologischen Forschung dominierende Perspektive auf *Moral Judgements*, indem sie über die bloßen individuellen Zuschreibungen von moralisch ‚richtig‘ oder ‚falsch‘ hinausgehen und vielmehr die Frage ins Zentrum rücken, welche Person es *gut* wäre zu sein (Hitlin & Stephen, 2010). Eng damit verbunden ist die Abgrenzung der *Ethics of Doing* und *Ethics of Being* (Abend, 2012). So haben zumindest viele Menschen eine bestimmte Vorstellung davon, was sie im Leben erreichen wollen und wer (welche Person) sie sein wollen. Um diesen Zielvorstellungen zu entsprechen, spielen von den Akteuren internalisierte und akzeptierte MVGs nicht nur unmittelbar im Kontext konkreter Handlungsentscheidungen eine zentrale Rolle, vielmehr besitzen sie auch eine identitätsstiftende Funktion. Erst ihre Einhaltung gewährleistet es, die Person zu sein, die man gerne sein möchte.

⁷Moralische Realisten gehen von der Existenz universalistischer MVGs aus. Demgegenüber lehnen moralische Skeptiker die Annahme einer solchen ab und vertreten ein deskriptives Moralitätsverständnis.

Um die zurecht kritisierte Unterkomplexität *dünnere* Moralitätskonzepte zu überwinden und die skizzierte Differenzierung in *Ethics of Doing* und *Ethics of Being* angemessen zu berücksichtigen, planen wir, ALLMOR eine zweidimensionale Struktur zugrunde zu legen: Eine Skala soll handlungsorientierte (HMor) und eine zweite Skala seinsorientierte (SMor) allgemeine Moralvorstellungen abbilden. Konkret bezieht sich HMor darauf, in welchem Maße unterschiedliche MVGs handlungsleitend wirken, während mit SMor das persönliche *Moral Make-Up* der Befragten erfasst werden soll. Zu den Details der Skalenkonstruktion siehe Abschnitt 4.1.

Abschließend gilt es noch die zuvor aufgeworfene Frage zu klären, welche Verhaltensgrundsätze vor dem Hintergrund des geplanten Einsatzbereichs der Skala als ‘moralisch’ zu bewerten sind. Hierfür rekurren wir auf die maßgeblich von Jonathan Haidt (Graham et al., 2013; Haidt, 2012; Haidt & Joseph, 2004) entwickelte *Moral Foundations Theory* (MFT). Die MFT basiert auf einer stark systemisch ausgerichteten Moralitätsdefinition und geht von der Prämisse aus, dass das moralische Urteilen bzw. Handeln von Individuen auf der Grundlage von fünf (bzw. sechs)⁸ *Moral Foundations* (MF) erfolgt. Diese bilden die *Moral Domain* und setzen sich zusammen aus: (i) *Care* (sich um andere Personen kümmern bzw. diese beschützen und wertschätzen), (ii) *Fairness* (die Ablehnung von Schummeln und Betrügen bzw. der unfairen Behandlung anderer Personen; Gerechtigkeit walten lassen;), (iii) *Loyalty* (das Einstehen für die eigene Familie, Gruppe bzw. das eigene Land; die Pflichten, die mit der Zugehörigkeit einhergehen, erfüllen), (iv) *Authority* (die gegebenen Traditionen und rechtlich legitimierte Autoritäten respektieren; Gehorsam zeigen; seine formalen Pflichten erfüllen) und (v) *Sanctity* (Abscheu vor widerlichen Dingen). Diese Ausgestaltung der *Moral Domain* stellt eine Erweiterung der *Moral Domain Theory* (DMT) von Turiel (1983), der diese definitorisch auf “prescriptive judgements of justice, rights, and welfare pertaining to how people ought to relate to each other” (Turiel, 1983, S. 3) beschränkt. Alle Bereiche, die nicht *Justice*, *Rights* und *Welfare* umfassen, werden von Turiel als ‘nicht-moralisch’ angesehen und vornehmlich dem Bereich der sozialen Konventionen zugeordnet. Der FMT liegt im Vergleich zur DMT somit ein breiteres Verständnis davon zugrunde, welche Regeln bzw. Verhaltensgrundsätze als moralisch zu bewerten sind. Allerdings räumt sie *Care* und *Fairness* als jenen beiden MFs, die mit der Definition der *Moral Domain* nach Turiel kongruieren, zumindest implizit eine gewisse konzeptionelle Priorität ein. Zur inhaltlichen Ausgestaltung der Items im Rahmen der Skalenkonstruktion von ALLMOR liegen mit den fünf MF nun schließlich auch konkrete theoretische Anhaltspunkte darüber vor, welche aus dem Universum aller Verhaltensgrundsätze als moralisch gelten, d.h. MVGs repräsentieren und sich zur Abbildung allgemeiner handlungs- und seinsorientierter Moralvorstellungen eignen.

⁸Haidt (2012) erweitert die ursprünglich fünf *Moral Foundations* um *Liberty*. Diese bleibt im vorliegenden Beitrag allerdings unberücksichtigt. Aus Platzgründen ist es weiterhin nicht möglich (und an dieser Stelle auch nicht notwendig), näher auf die MFT einzugehen. Interessierte Leserinnen und Leser seien hierfür insbesondere auf Haidt (2012) verwiesen.

3 *Deviant Morality Scales* – Eine knappe kritische Reflexion

Wie bereits in Abschnitt 1 erwähnt, genießen *Deviant Morality Scales* (DMS) in der empirischen Kriminalsoziologie bzw. Kriminologie gegenwärtig breite Popularität. Dies ist in erster Linie dem Umstand geschuldet, dass sich das Design der PADS+-Studie⁹ (inklusive der darin enthaltenen Messkonzepte) in den letzten Jahren als standardmäßiges methodisch-methodologisches Framework zur empirischen Prüfung von (SAT-basierten) Hypothesen über den Einfluss der allgemeinen Moralvorstellungen von Akteuren auf deren kriminalitätsbezogene Handlungsentscheidungen etabliert hat. Allgemein zielen DMS darauf ab, allgemeine Moralvorstellungen über das Ausmaß der Ablehnung normativ geächteter Handlungen zu erfassen. Konkret enthält etwa die PADS+ zugrunde liegende SAT-MDS insgesamt 16 Items (Wikström et al., 2012, S. 134, siehe dazu auch Abschnitt 4.3), die jeweils erfassen, für wie *falsch* es die Befragten erachten, wenn jemand ihres Alters eine bestimmte Normverletzung begeht. Hierbei handelt es sich beinahe ausschließlich um Handlungen, die gegen kodifizierte Normen (z.B. Schulordnung, Verordnungen, Jugendschutzbestimmungen, Straßenverkehrsordnung, Strafrecht) verstoßen. An dieser Stelle offenbart sich eine Reihe konzeptioneller Limitierungen von MDS, die ihre Eignung zur adäquaten Abbildung individueller Moralvorstellungen beträchtlich in Zweifel ziehen.

(i) Da DMS *per definitionem* lediglich ein hochselektives Set an Handlungen enthalten, werden sie dem in Abschnitt 2 dargelegten universalistischen Verständnis von allgemeinen Moralvorstellungen nicht gerecht. (ii) Ein ausschließlich handlungsbezogenes Messkonzept ignoriert ferner die Perspektive der *ethics of being* und erlaubt im besten Fall die Abbildung eines *dünnen* Konzepts von Moralität. (iii) Von einem Messkonzept, das den Anspruch erhebt allgemeine Moralvorstellungen abzubilden, muss erwartet werden, dass nicht nur das Ausmaß der Ablehnung ‘unmoralischer’ Handlungen Berücksichtigung findet, sondern in gleichem Maße auch die Intensität der Befürwortung ‘moralischer’ Handlungen. Dies trifft auf DMS nicht zu. (iv) Es ist anzunehmen, dass die Ablehnung von Handlungen, die gegen kodifizierte (Rechts-)Normen verstoßen, stärker vom individuellen Ausmaß der Internalisierung und Übernahme exakt dieser Normen als handlungsleitende Prinzipien determiniert wird als von allgemeinen Moralvorstellungen (siehe dazu auch die entsprechende Argumentation in Abschnitt 2). Daraus folgt, dass DMS vielmehr den Konstruktkomplex *Normtreue, -internalisierung, -orientierung* bzw. *-bindung* abzubilden scheinen als allgemeine Moralvorstellungen.¹⁰ (v) Die mit der Verwendung von DMS implizit getroffene Annahme, dass das Ausmaß der Ablehnung bestimmter Handlungen einen Einfluss darauf hat, mit welcher Wahrscheinlichkeit/Häufigkeit mehr oder weniger dieselben Handlungen (dies gilt insbesondere bei der Verwendung von Versatilitäts- und über mehrere Delikte kumulierte Häufigkeitsska-

⁹PADS steht für *Peterborough Adolescent and Young Adult Development Study*. Hierbei handelt es sich um eine seit 2002 andauernde Panelerhebung von mehr als 700 Jugendlichen (nun jungen Erwachsenen) in der Stadt Peterborough (UK). Die Studie wird von Per-Olof Wikström geleitet und ist an der University of Cambridge angesiedelt. Für weitere Details sei auf <http://www.pads.ac.uk> verwiesen.

¹⁰An dieser Stelle muss angemerkt werden, dass es sich hierbei um eine theoretisch begründete Annahme handelt, die einer empirischen Prüfung bedarf.

len) tatsächlich auch begangen werden, erscheint trivial und verfehlt das eigentliche Ziel, die Wirkung von allgemeinen Moralvorstellungen auf kriminalitätsbezogene Handlungsentscheidungen empirisch prüfen zu wollen. Vielmehr stellt sich die Frage, in welchem Maße es sich hierbei um eine konzeptionelle Tautologie handelt. (vi) Da sich die in DMS enthaltenen Normverstöße entlang unterschiedlicher Dimensionen ausdifferenzieren (z.B. entlang der strafrechtlichen Relevanz, des sozialen Kontexts, Selbst- bzw. Fremdschädigung), weisen sie in aller Regel eine mehrfaktorielle Struktur und somit keine günstigen psychometrischen Eigenschaften auf.

Vor dem Hintergrund der skizzierten Defizite erscheint der Bedarf an einem alternativen Messkonzept für allgemeine Moralvorstellungen offensichtlich. In der Folge wird versucht, ein solches auf der in Abschnitt 2 erarbeiteten Grundlage zu entwickeln.

4 Methodische Vorgehensweise

4.1 Skalenkonstruktion

Entsprechend den Ausführungen in Abschnitt 2 soll das Instrument ALLMOR aus zwei Skalen zur Abbildung individueller Moralvorstellungen bestehen, einer handlungsorientierten (HMor) und einer seinsorientierten (SMor) Skala. Da es mit dem Ziel konstruiert wird, als ‚typisches‘ Survey-Instrument im Rahmen der kriminalsoziologischen bzw. kriminologischen Forschung eingesetzt zu werden, ist eine Beschränkung der Anzahl an Items auf etwa fünf pro Skala geplant. Damit soll die praktische Einsetzbarkeit des Instruments vor dem Hintergrund limitierter Platzressourcen in Erhebungsinstrumenten zu selbstberichteter Delinquenz sichergestellt werden. Ferner ergibt sich daraus auch die Bezugspopulation von ALLMOR, die mit der erwachsenen Allgemeinbevölkerung westlicher Gegenwartsgesellschaften auch eine der idealtypischen Zielpopulationen von Surveys zu selbstberichteter Delinquenz im deutschsprachigen Forschungsraum repräsentiert.

Die konkrete inhaltliche Ausgestaltung¹¹ der moralischen Verhaltensgrundsätze im Rahmen der Itemformulierung orientiert sich für beide Skalen an den fünf *Moral Foundations* (MF) der *Moral Foundations Theory* (Abschnitt 2): (i) *Care* (C), (ii) *Fairness* (F), (iii) *Loyalty* (L), (iv) *Authority* (A) und (v) *Sanctity* (S). Aufgrund der übergeordneten Relevanz von *Care* und *Fairness* werden die beiden Dimensionen priorisiert, d.h. es werden mehr Items konzipiert als für die restlichen drei MF. Um potentiellen Dimensionalitätsproblemen entgegenzuwirken, werden die MF auch mit unterschiedlichem Gewicht auf die beiden Skalen verteilt. Während HMor stärker mit Items zu *Fairness* und *Loyalty* besetzt ist, sind in SMor Items zu *Authority* und *Sanctity* überrepräsentiert. *Care* ist in beiden Skalen annähernd in gleichem Maße vertreten. Weiterhin werden in HMor zum Zwecke der diskriminanten Validitätsprüfung vier Items (die Items i3_h, i4_h, i9_h und i18_h in Tabelle 8) integriert, die das Konstrukt *Egozentrismus* abbilden. Analog dazu enthält SMor drei Items (die Items i13_s, i15_s und i17_s in Tabelle 9), die ähnlich wie DMS die

¹¹Wir danken Elena Fattorelli, Theresa Fruß, Michaela Kotscherwa, Rouven Link und Kerstin Schieber, allesamt Teilnehmerinnen und Teilnehmer des Forschungspraktikums „Werte und Kriminalität“ im Bachelorstudiengang Soziologie an der Universität Eichstätt-Ingolstadt im Studienjahr 2015/16, für ihren bedeutsamen Beitrag zur Itementwicklung.

Akzeptanz strafbarer Handlungen abbilden. Um eine angemessene Selektion der Items mit den besten Messeigenschaften zu ermöglichen, werden pro Skala anfänglich etwa 20 Items entwickelt (siehe die Tabellen 8 und 9 im Anhang).

Im Zuge der Skalenkonstruktion von HMor rekurren wir zur optimalen Ausdifferenzierung der Itemschwierigkeiten außerdem auf das Stufenmodell der moralischen Urteilsentwicklung von Kohlberg (1974, 1996).¹² Dieses strukturiert den Prozess der Ausbildung des Moralbewusstseins prinzipiell in drei Ebenen mit je zwei Stufen (präkonventionell mit den Stufen 1-2, konventionell mit den Stufen 3-4, postkonventionell mit den Stufen 5-6), die von den Individuen sequentiell durchlaufen werden. Nicht alle Individuen erreichen allerdings tatsächlich auch höhere Stufen, insbesondere nicht jene der postkonventionellen Ebene. Konkret orientieren wir uns am *Prinzip der Erweiterung der sozialen Perspektive*. Demzufolge extendiert die Wahrnehmung des von Individuen als urteilsrelevant erachteten sozialen Kontexts im Zuge der Entwicklung des Moralbewusstseins von einer anfänglich rein egozentrischen Perspektive auf der präkonventionellen Ebene hin zu einer universalistisch geprägten Perspektive auf der postkonventionellen Ebene. Das Level der wahrgenommenen sozialen Perspektive kongruiert folglich mit der Ebene der Moralentwicklung, auf der sich ein Individuum gerade befindet. Daraus lässt sich ableiten, dass Variation in der sozialen Perspektive Variation in den Itemschwierigkeiten induziert. Da Items mit beispielsweise universalistischem Kontext nur von jenen Befragten als (sehr) zutreffend beurteilt werden können, die auch tatsächlich die postkonventionelle Ebene erreicht haben, weist ein solches Item eine höhere Itemschwierigkeit auf als ein Item, dessen soziale Perspektive jener der konventionellen Ebene entspricht. Darauf aufbauend werden die Items von HMor entlang von drei Kategorien formuliert: Sie beziehen sich inhaltlich auf (I) die egozentrische Perspektive (präkonventionell), (II) die (un-)mittelbare soziale und gesellschaftliche Perspektive (konventionell) sowie (III) die universelle Perspektive (postkonventionell; siehe dazu auch Tabelle 8).

Beiden Skalen wird eine fünfstufige Ratingskala mit den Ausprägungen 1 (trifft gar nicht zu), 2 (trifft eher nicht zu), 3 (trifft teilweise zu), 4 (trifft eher zu), 5 (trifft völlig zu) zugrunde gelegt. Zudem werden neben positiv formulierten (in Richtung der Zieldimension) Items jeweils auch einige invers formulierte Items entwickelt und getestet.

4.2 Datengrundlage

Das der analytischen Konstruktion des Messinstruments zugrunde liegende Design sieht zunächst die Erhebung von Daten einer Entwicklungsstichprobe (in der Folge als *Sample 1* bezeichnet) vor, gefolgt von einer zweiten unabhängigen Validierungsstichprobe (*Sample 2*). Die Datenerhebungsprozeduren für beide Stichproben werden nachfolgend beschrieben.

Die *Sample 1* zugrundeliegende Grundgesamtheit wurde definiert als alle in Ingolstadt (Bayern) mit Hauptwohnsitz gemeldeten, volljährigen Personen mit deutscher Staats-

¹²Während Kohlberg zunächst moralisches Urteilen von moralischem Handeln trennt und das Stufenmodell auf Ersteres bezieht, postuliert er später, dass insbesondere auf hohen Entwicklungsstufen moralisches Urteilen auch moralisches Handeln zur Folge hat.

angehörigkeit.¹³ Das Melderegister der Stadt bildete die Auswahlgrundlage (Stichtag: 09. März 2016) für die Ziehung einer vom Bürgerservice Ingolstadt durchgeführten einfachen Zufallsauswahl der Größe $n_1 = 2.000$. Die Erhebung wurde unter Berücksichtigung zentraler Aspekte der *tailored design method* (TDM; Dillman et al., 2014) als Querschnitt im postalischen Mode administriert. Die Feldphase dauerte von 17. März bis 21. Mai 2016. Zur Optimierung der Rücklaufquote wurde am 21. April 2016 ein einmaliges Erinnerungsschreiben versendet. Von allen Personen der Bruttostichprobe mussten 60 (3,0 % der Bruttostichprobe) als neutrale Ausfälle (Fragebogen postalisch nicht zustellbar) klassifiziert werden. Insgesamt wurden 598 Fragebögen retourniert. Davon mussten 43 Fragebögen aussortiert werden, da sie entweder völlig unausgefüllt waren oder gemäß den AAPOR-Konventionen (AAPOR, 2016, S. 29) einen inakzeptabel hohen Anteil an fehlenden Werten (> 50 %) aufwiesen (*break-offs*). Die vorliegende realisierte Nettostichprobe umfasst schließlich $n_{1r} = 555$ Befragte. Dies entspricht einer Rücklaufquote von 27,8 % (nach RR6 gemäß AAPOR, 2016, S. 16).

Die *Sample 2* zugrunde liegende Grundgesamtheit setzt sich aus allen volljährigen deutschen Staatsangehörigen bis zu einem maximalen Alter von 65 Jahren zusammen, deren Hauptwohnsitz (und einziger Wohnsitz) in Frankfurt am Main liegt. Wiederum bildete das Melderegister der Stadt die Auswahlgrundlage (Stichtag: 28. Februar 2019) für die Ziehung einer nach dem Geschlecht und spezifischen Altersgruppen (18-32 Jahre, 33-49 Jahre, 50-65 Jahre) disproportional stratifizierten Zufallsauswahl der Größe $n_2 = 3.000$ (jeweils 500 Männer und Frauen). Analog zu *Sample 1* wurde die Erhebung nach den Grundsätzen der TDM konzipiert und als Querschnitt in postalischer Form realisiert. Die Feldphase begann am 18. März 2019 – eine Woche vor der postalischen Versendung der Fragebögen – mit der Aussendung einer Vorabinformation und dauerte bis 16. Juni 2019. Ein einmaliges Erinnerungsschreiben wurde am 06. Mai 2019 verschickt. Die Bruttostichprobe enthielt schließlich 13 neutrale Ausfälle (0,4 %). Von den insgesamt 464 retournierten Fragebögen mussten wiederum 15 wegen einer Überschreitung des festgelegten Schwellenwerts für den Anteil an fehlenden Werten (> 50 %) aussortiert werden, so dass sich die realisierte Nettostichprobe aus $n_{2r} = 449$ Befragten zusammensetzt. Dies entspricht einer Rücklaufquote (RR6) von 15,0 %.¹⁴

4.3 Operationalisierung

Im Rahmen der Prüfung der Kriteriumsvalidität und des Einflusses sozial erwünschten Antwortverhaltens auf ALLMOR (siehe Abschnitt 4.4) spielen noch eine Reihe von anderen Merkmalen eine bedeutende Rolle, deren Operationalisierung in der Folge erläutert wird.

Zunächst sollen die beiden verfügbaren Merkmale zu selbstberichteter Delinquenz,

¹³Die Einschränkung der Grundgesamtheit auf Personen mit deutscher Staatsangehörigkeit ist dem Umstand geschuldet, dass der Fragebogen aus Kostengründen nur deutschsprachig zur Verfügung gestellt werden konnte. Ziel war es, die Instrumententwicklung nicht durch eine reduzierte Messqualität der Indikatoren zu beeinträchtigen, die auf Probleme im Sprachverständnis zurückzuführen ist.

¹⁴Bei Interesse können Detailinformationen zur Komposition der beiden Stichproben von den Autoren angefordert werden.

die jeweils ‚typische‘ Vergehen der erwachsenen Allgemeinbevölkerung („Kriminalität im Alltag“ Eifler, 2009) darstellen, beschrieben werden. Die Daten der Entwicklungsstichprobe (*Sample 1*) enthalten die Einjahresprävalenz zu *Schwarzfahren mit öffentlichen Verkehrsmitteln* (diese liegt bei 3,1 %; SE: 0,73), jene der Validierungsstichprobe (*Sample 2*) die Lebenszeitinzidenz zu *Trunkenheit am Steuer*. Für letztere wurde konkret erfragt, wie häufig die Befragten in ihrem Leben bereits mit dem Auto gefahren sind, obwohl sie die zulässige Promillegrenze mit Sicherheit überschritten hatten. Das Merkmal wurde ordinal erfasst (Ausprägungen: noch nie / einmal / 2-5 mal / 6-10 mal / 11-20 mal / mehr als 20 mal) und zum Zwecke der Integration in lineare Strukturgleichungsmodelle (siehe Abschnitt 4.4 als endogene Variable dichotomisiert (noch nie / mindestens einmal). Die in die Modelle eingehende Information entspricht somit der Lebenszeitprävalenz (50,8 %; SE: 2,4).

Zur Prüfung der diskriminanten Validität und komparativen Evaluation der Erklärungskraft von ALLMOR (siehe Abschnitt 4.4) wurde in beiden Erhebungen auch eine DMS implementiert. Hierbei handelt es sich um eine für die erwachsene Allgemeinpopulation adaptierte 7-Item Kurzversion der 16-Item SAT-DMS (Wikström et al., 2012, S. 134). Die verwendeten Items sind in Tabelle 7 im Anhang ausgewiesen. Entfallen sind jene Items, deren Inhalt sich explizit auf Jugendliche bezieht und für Erwachsene nicht von Bedeutung ist (z.B. den Bleistift eines Klassenkameraden stehlen, die Hausaufgaben nicht machen, der Schule ohne Entschuldigung fernbleiben). Die zugrunde gelegte vierstufige Ratingskala von 1 (sehr falsch) bis 4 (überhaupt nicht falsch) wurde ohne Änderung übernommen. Explorative Hauptkomponentenanalysen (PCA; zur Vorgehensweise siehe Abschnitt 4.4) indizieren für beide Stichproben eine zweifaktorielle Struktur. Der Bruch in den Items erfolgt entlang der Deliktsschwere (siehe dazu auch Abschnitt 3). Während in *Sample 2* nur die beiden aus strafrechtlicher Sicht schwerwiegendsten Normverstöße (Items d5 und d7) einen zweiten Faktor bilden, gehören in *Sample 1* auch die beiden etwas geringfügigeren Delikte (Items d2 und d4) diesem Faktor an. Diese besitzen allerdings in beiden Stichproben keine klare Einfachstruktur.¹⁵ Da die eindimensionale Behandlung der Skala der Anwendungspraxis entspricht (und ALLMOR gegen diese getestet werden soll), werden wir trotz dieser faktoranalytischen Befundlage die sieben Items als Indikatoren eines latenten Faktors begreifen. Für eine solche eindimensionale Lösung variieren die PCA-basierten Faktorladungen der Items zwischen 0,50 und 0,70 (*Sample 1*) bzw. 0,29 und 0,77 (*Sample 2*), der Eigenwert des Faktors beträgt 2,66 (er bindet 38,0 % der Gesamtvarianz aller Items; *Sample 1*) bzw. 2,71 (38,7 %; *Sample 2*). Da die Items das latente Konstrukt nicht eindimensional repräsentieren, wird auf die Reliabilitätsprüfung verzichtet.

Zur Messung *sozial erwünschten Antwortverhaltens* wurde in *Sample 1* die 6-Item ‚Kurzskala Soziale Erwünschtheit-Gamma‘ (KSE-G; Kemper et al., 2012) und in *Sample 2* die 10-Item Skala der Dimension ‚Fremdtäuschung‘ (auch als ‚Eindrucksmanagement‘ bezeichnet) des Instruments von Musch et al. (2002) verwendet. Beide Instrumente basieren auf dem zweidimensionalen Konzept sozialer Erwünschtheit nach Paulhus (1984, 1986, 1991, 2002), das eine Differenzierung in Selbst- und Fremdtäuschung vorsieht.

¹⁵Bei Interesse können die detaillierten Ergebnisse von den Autoren angefordert werden.

Während die aus dem Instrument von Musch et al. (2002) ausgewählte Skala ausschließlich auf die Messung von Fremdtäuschung abzielt, bildet der KSE-G den *Gamma-Faktor* ab. Dieser beinhaltet sowohl Elemente der Fremd- als auch der Selbsttäuschung, deren Gemeinsamkeit der inhaltliche Bezug auf *Communion* (z.B. Gemeinsinn, Verbundenheit, Kooperation) ist und die – so die Vermutung – im Zuge der Beantwortung von Items zu allgemeinen Moralvorstellungen und selbstberichteter Delinquenz zulasten Dritter bzw. der Allgemeinheit aktiviert werden.¹⁶ Mit der Verwendung der beiden unterschiedlichen Instrumente für die Entwicklungs- und Validierungserhebung war das Ziel verbunden, ein umfassendes Abbild des Einflusses sozial erwünschten Antwortverhaltens auf ALLMOR vorzulegen, das nicht von der Messqualität eines einzigen Instruments abhängt.

4.4 Analytische Vorgehensweise

Die analytische Arbeit folgt einer mehrstufigen Struktur: Am Beginn steht (i) die Skalenentwicklung inklusive Reliabilitätsanalyse, im Anschluss daran erfolgt (ii) die Validitätsprüfung inklusive (iii) einer Evaluation, wie sensitiv die entwickelten ALLMOR-Skalen auf sozial erwünschtes Antwortverhalten reagieren.

(i) Im Rahmen der Skalenentwicklung werden die Itembatterien für handlungs- (HMor) und seinsorientierte Moralvorstellungen (SMor) zum Zwecke der explorativen Dimensionalitätsüberprüfung getrennt voneinander mittels Hauptkomponentenanalysen (*principal component analysis*; PCA; siehe z.B. einführend Wolff & Bacher, 2010) analytisch bearbeitet. Zur Bestimmung der relevanten Komponentenanzahl wird das Paralleltestverfahren (Horn, 1965) zum Einsatz gebracht (hierzu wird das Stata add-on `paran` verwendet; Dinno, 2009). Dieses sieht eine Testung der vorliegenden empirischen Lösung gegen den Mittelwert der Lösungen aus einer Reihe von generierten Zufallsdaten vor, die in Stichprobengröße und Itemanzahl exakt den empirischen Daten entsprechen. Gemäß Paralleltestkriterium werden jene Komponenten extrahiert, deren Eigenwerte über den gemittelten Eigenwerten der entsprechenden Komponenten aus den Zufallsdaten liegen. Zur Optimierung der Itemzuordnung zu den extrahierten Komponenten wird bei mehrdimensionalen Lösungen zudem das Varimax-Verfahren als orthogonale Methode der Achsenrotation angewendet. Die Vorgehensweise erscheint gerechtfertigt, da den jeweiligen Itembatterien konstruktionsbedingt a priori Eindimensionalität unterstellt wird und im Falle einer Abweichung der dimensional Struktur eine klare Abgrenzung zu den anderen latenten Faktoren angestrebt wird. Ziel der PCA ist schließlich die Selektion jener Items, die eindimensional mit hinreichend hoher Faktorladung (Kriterium: $|\lambda| > 0,5$) den jeweiligen latenten Faktor (HMor bzw. SMor) repräsentieren.

Im Anschluss werden die beiden entwickelten Skalen einer Reliabilitätsanalyse unterzogen. Konkret wird die interne Konsistenz auf Basis des Koeffizienten *Omega* (McDonald, 1999) ermittelt. Im Gegensatz zum weitverbreiteten Koeffizienten *Alpha* (Cronbach, 1951) setzt *Omega* nicht die empirisch selten erfüllte Annahme essentieller tau-Äquivalenz (Gleichheit der Faktorladungen aller Indikatoren eines latenten Konstrukts)

¹⁶Im Gegensatz dazu repräsentiert der *Alpha-Faktor* Elemente der Fremd- und Selbsttäuschung, die sich auf *Agency* (z.B. Selbstbehauptung, Selbstentfaltung, Abgrenzung) beziehen. Zu weiteren Details siehe Paulhus (2002).

voraus, sondern hat das weniger restriktive kongenerische Messmodell zur Grundlage. Demzufolge konnten Zinbarg et al. (2005) formal sowie Trizano-Hermosilla und Alvarado (2016) simulationsbasiert nachweisen, dass *Omega* bei der Schätzung der internen Konsistenz *Alpha* klar überlegen ist, wenn eine Verletzung der tau-Äquivalenzannahme vorliegt, in jedem Fall aber niemals schlechter abschneidet.¹⁷ Aus Gründen der Zweckmäßigkeit wird zusätzlich allerdings auch *Alpha* berichtet. Da es sich bei *Alpha* und *Omega* um Stichprobenparameter handelt, die mit Unsicherheit behaftet sind (dies wird von der Mehrheit an empirischen Studien ignoriert; siehe dazu Dunn et al., 2014), werden für beide Koeffizienten auch bootstrapbasierte Konfidenzintervalle (siehe z.B. Dunn et al., 2014; Padilla & Divers, 2013; Zhang & Yuan, 2016) basierend auf $B = 1.000$ Bootstrap-Samples ausgewiesen. Diese werden mit dem R-Paket MBESS (Kelley, 2019) ermittelt.¹⁸

(ii) Im ersten Schritt der Validitätsprüfung erfolgt die Evaluation der *faktoriellen Validität* der Skalen HMor und SMor mittels konfirmatorischer Faktorenmodelle (*confirmatory factor analysis*; CFA; z.B. Brown, 2006; Reinecke, 2014). Zunächst wird die Passung der beiden Skalen an die empirischen Daten in getrennten Modellen (M1 und M2; für eine zusammenfassende Übersicht zur Spezifikation aller im Anschluss geschätzten Modelle siehe Tabelle 1) evaluiert. In Anschluss daran erfolgt die Prüfung des Zwei-Faktorenmodells (M3), inklusive der Ermittlung der entsprechenden Faktorkorrelation. Da beide Faktoren dasselbe latente Konstrukt repräsentieren und sich lediglich auf unterschiedliche Aspekte (Handeln und Sein) davon beziehen, ist eine hohe Faktorkorrelation ($> 0,5$) zu erwarten. In diesem Zusammenhang soll ferner die Angemessenheit des Ein-Faktorenmodells für beide Skalen (M4; die latente Variable G in Tabelle 1 repräsentiert diesen Faktor) geprüft werden. Dies hat vorwiegend skalierungspragmatische Gründe: Aus messtheoretischer Sicht erscheint es plausibel, HMor und SMor als Faktoren erster Ordnung des zugrunde liegenden latenten Konstrukts ‚allgemeine Moralvorstellungen‘ als Faktor höherer Ordnung aufzufassen. Allerdings ist die empirische Umsetzung eines solchen Faktorenmodells (siehe einfürend z.B. Brown, 2006; Reinecke, 2014) im gegebenen Fall nicht möglich, da zur Identifikation eines Faktors zweiter Ordnung zumindest drei latente Konstrukte auf der ersten latenten Ebene vorhanden sein müssen. Dies wirft in der Folge die Frage auf, wie mit den beiden Skalen umgegangen werden soll. Prinzipiell stehen zwei Optionen zur Auswahl: (i) Beide Skalen werden als eigenständig angesehen und analytisch als getrennte Messungen behandelt oder (ii) sie werden zu einem Faktor fusioniert. Insgesamt sprechen zentrale Argumente für die zweite Option. Zunächst trägt die Verwendung einer gemeinsamen Skala zur Modellsparsamkeit bei und erlaubt eine einfache Interpretation des Einflusses allgemeiner Moralvorstellungen auf interessierende soziale Phänomene, wie etwa delinquentes Handeln. Weiterhin entspricht es der Idee, dass beide Skalen dasselbe latente Konstrukt repräsentieren und eben nicht unabhängig voneinander sind. Vor diesem Hintergrund erscheint es somit zweckmäßig, das Ein-Faktormodell als ernsthafte Skalierungsvariante empirisch zu prüfen.

¹⁷Für substantielle Kritik an *Alpha* siehe ferner z.B. Dunn et al. (2014), Raykov (1998), Sijtsma (2009) sowie Yang und Green (2011).

¹⁸Für diese Aufgabe steht auch das R-Paket `coefficientalpha` (Zhang & Yuan, 2016) zur Verfügung, die Resultate sind im vorliegenden Fall mehr oder weniger identisch.

Im zweiten Schritt wird die *Kriteriumsvalidität* der beiden Skalen evaluiert. In diesem Zusammenhang bietet es sich selbstverständlich an, delinquentes Handeln (C in M5 bis M12) als Außenkriterium heranzuziehen und zu prüfen, ob der aufgrund der theoretischen Annahmen erwartbare negative Einfluss – die ausgeprägte Übernahme allgemeiner Moralvorstellungen reduziert die Wahrscheinlichkeit, dass Akteure delinquente Handlungen überhaupt als Gelegenheit wahrnehmen bzw. sich dafür entscheiden (siehe dazu Abschnitt 1) – empirisch auch tatsächlich beobachtet werden kann. Zu diesem Zweck werden M1 bis M4 zur Schätzung des entsprechenden Effektparameters jeweils um eine Strukturgleichung erweitert, in der C als lineare Funktion von HMor und SMor bzw. *G* spezifiziert wird (M5 bzw. M8).

Den dritten Schritt der Validitätsprüfung stellt schließlich die Evaluation der *diskriminanten Validität* zur SAT-DMS (*D* in M9 bis M12) dar. Hierfür werden die Modelle M7 und M8 um das entsprechende Messmodell erweitert, zwischen allen endogenen latenten Variablen werden ungerichtete Pfade (Kovarianz- bzw. Korrelationsparameter) spezifiziert (M9 und M10). Wie bereits in Abschnitt 3 ausgeführt wurde, ist davon auszugehen, dass die SAT-DMS keine valide Messung allgemeiner Moralvorstellungen repräsentiert, sondern vielmehr das Ausmaß an Normtreue bzw. -bindung abbildet. Da die Grundlage hierfür ebenfalls internalisierte Verhaltensnormen darstellen, ist zu erwarten, dass die Konstrukte zwar positiv, nicht aber substantiell ($< 0,3$) miteinander korrelieren. Weiterhin soll geprüft werden, ob der erwartete negative Effekt von HMor und SMor bzw. *G* auch bestehen bleibt (sich nicht substantiell verringert), wenn das Modell um die SAT-DMS erweitert wird. In diesem Zusammenhang wird ebenfalls ein Vergleich der Erklärungskraft der Konstrukte (HMor & SMor bzw. *G* vs. SAT-DMS) über deren standardisierte Effektstärken vorgenommen.

(iii) Der finale Analyseschritt umfasst die Prüfung der ALLMOR-Skalen im Hinblick auf ihre Anfälligkeit für sozial erwünschtes Antwortverhalten. Dies erfordert die Erweiterung der Modelle M9 und M10 um die latente exogene Variable *soziale Erwünschtheit* (*E* in M11 und M12). Von Interesse sind einerseits die Korrelationen von HMor, SMor, *G* sowie SAT-DMS mit *E* und andererseits, ob sich durch die Aufnahme von *E* die Schätzer der Effektparameter der Moralitätskonstrukte substantiell verringern, da sozial erwünschtes Antwortverhalten eine gemeinsame Ursache (*common cause*) systematischer Antwortmuster in **h**, **s** bzw. **d** (Indikatorensatz für HMor, SMor, und SAT-DMS) und C repräsentiert.

Die Parameterschätzung erfolgt für alle Modelle auf Basis des *Weighted Least Squares* (WLS) Schätzers (Browne, 1982, 1984), da für diesen keine Annahmen über Schiefe und Kurtosis – weder für die Verteilungen der Indikatoren noch für jene der latenten Variablen – getroffen werden müssen (Reinecke, 2014, S. 105). Die Wahl der Schätzers muss vor dem Hintergrund gesehen werden, dass bei gegebener Datenlage die Erfüllung der multivariaten Normalverteilungsannahme des konventionellen *Maximum Likelihood* Schätzers nicht erwartet werden kann. Um die Metrik der latenten Variablen in den Messmodellen festzulegen, wird deren Varianz jeweils auf 1 fixiert. Da es sich weiterhin bei C um dichotome abhängige Variable (konkret um Prävalenzen; siehe Abschnitt 4.3) handelt, werden die Befunde im Rahmen eines Robustheitstests noch über die Schätzung von generalisierten Strukturgleichungsmodellen mit Logit-Link abgesichert (dies betrifft

M5 bis M12).¹⁹ Zur Evaluation des Modellfits greifen wir in Anlehnung an die Empfehlung von Kline (2005) auf den RMSEA (Schwellenwert: $RMSEA_S = 0,08$; niedrigere Werte \rightarrow besserer Fit), den CFI ($CFI_S = 0,9$; höhere Werte \rightarrow besserer Fit) und den SRMR ($SRMR_S = 0,08$; niedrigere Werte \rightarrow besserer Fit) zurück. Liegen zumindest zwei der drei Fit Indizes über (bzw. unter) dem Schwellenwert, wird eine hinreichend gute Passung des Modells an die Daten unterstellt. Die Analysen werden mit `Stata SE 15.0`, `Mplus 7.11` und `R 4.0.0` realisiert.

5 Ergebnisse

Für die gesamte nachfolgende Ergebnisdarstellung gilt, dass die Befunde aus Entwicklungs- und Validierungsstichprobe für jeden einzelnen Analyseschritt parallel (vergleichend) dargestellt werden.

5.1 Skalenentwicklung

Die Ergebnisse der PCA für die ursprünglich 19 Items (siehe Tabelle 8) zur Entwicklung der handlungsorientierten Skala (HMor) sind in Tabelle 2 aufbereitet. Insgesamt weisen die Daten eine dreidimensionale Struktur auf. Auf der ersten Komponente (K1) laden insgesamt sechs Items ($i2_h, i5_h, i6_h, i13_h, i14_h, i19_h$), bei einem Eigenwertanteil der Komponente von 14,4 %. Wie sich zeigt, xc repräsentieren mit Ausnahme von $i14_h$ (*Fairness*) alle Items die *Moral Foundation* (MF) *Care* und sind hinsichtlich der sozialen Perspektive zumindest der konventionellen Ebene des Stufenmodells der Moralentwicklung zuzuordnen (siehe dazu auch die entsprechenden Informationen in Tabelle 8). Die zweite Komponente (K2) umfasst fünf Items ($i4_h, i7_h, i15_h, i16_h, i18_h$; Eigenwertanteil der Komponente: 11,3 %), die vornehmlich die egozentrische soziale Perspektive abbilden. Weiterhin gehören zwei der Items ($i4_h, i18_h$) den insgesamt vier aufgenommenen Egozentrismus-Items zur explorativen Prüfung der diskriminanten Validität an. Die Befunde legen somit den Schluss nahe, dass es sich bei K2 offensichtlich eher um eine Messung von Egozentrismus (Eigennützigkeit) handelt als um eine repräsentative Abbildung allgemeiner Moralvorstellungen. Auf der dritten Komponente (K3) laden insgesamt vier Items ($i3_h, i8_h, i10_h, i12_h$; Eigenwertanteil der Komponente: 0,9 %), die hinsichtlich der sozialen Perspektive über alle drei Ebenen variieren und – abgesehen von Item $i3_h$ (Egozentrismus) – inhaltlich die MF *Loyalty* abbilden. Die verbleibenden vier Items ($i1_h, i8_h, i10_h, i12_h$) verfügen über keine Einfachstruktur (sie laden auf keiner der extrahierten Komponenten mit einer Faktorladung in relevantem Maße) und werden bei der Interpretation der Komponenten ignoriert. In der Folge wird mit lediglich K1 weitergearbeitet, da (i) der Eigenwertanteil der Komponente mit Abstand am größten ist, (ii) die Komponente auf den meisten Items basiert und (iii) die überwiegende Mehrheit der Items mit *Care* eine der beiden zentralen MF abbildet. Aus inhaltlicher bzw. methodischer Sicht ist bedauerlich, dass die Komponente hinsichtlich der enthaltenen *Moral Foundations* nicht heterogener ausgestaltet ist und zum Zwecke der Identifikation

¹⁹Die Befunde des Robustheitstests werden aus Platzgründen nur berichtet, wenn substantielle Abweichungen in den Inferenzen im Vergleich zum linearen Modellen auftreten.

Modell	Spezifikation
M1	$H \rightarrow \mathbf{h}$
M2	$S \rightarrow \mathbf{s}$
M3	$\mathbf{h} \leftarrow H \leftrightarrow S \rightarrow \mathbf{s}$
M4	$\mathbf{h} \leftarrow G \rightarrow \mathbf{s}$
M5	$\mathbf{h} \leftarrow H \rightarrow C$
M6	$\mathbf{s} \leftarrow S \rightarrow C$
M7	$\mathbf{h} \leftarrow H \rightarrow C \leftarrow S \rightarrow \mathbf{s}$
M8	$\{\mathbf{h}, \mathbf{s}\} \leftarrow G \rightarrow C$
M9	$\{\mathbf{h} \leftarrow H, \mathbf{s} \leftarrow S\} \rightarrow C \leftarrow D \rightarrow \mathbf{d}$
M10	$\{\mathbf{h}, \mathbf{s}\} \leftarrow G \rightarrow C \leftarrow D \rightarrow \mathbf{d}$
M11	$\{\mathbf{h} \leftarrow H, \mathbf{s} \leftarrow S\} \rightarrow C \leftarrow \{D \rightarrow \mathbf{d}, E \rightarrow \mathbf{e}\}$
M12	$\{\mathbf{h}, \mathbf{s}\} \leftarrow G \rightarrow C \leftarrow \{D \rightarrow \mathbf{d}, E \rightarrow \mathbf{e}\}$

Anmerkungen: \mathbf{h} ... Indikatorenset der latenten Variable H (dies gilt in gleichem Maße für \mathbf{s} , \mathbf{d} und \mathbf{e}); H ... HMor, S ... SMor, G ... Gesamtfaktor allgemeine Moralvorstellungen; C ... delinquentes Handeln, D ... SAT-DMS, E ... soziale Erwünschtheit; latente Variable sind *kursiv* gesetzt; ab M3 werden sämtliche Korrelationen zwischen den endogenen Strukturmerkmalen S , H , G , D und E spezifiziert, auf eine explizite Abbildung wird zum Zwecke der Übersichtlichkeit verzichtet

Tabelle 1: Geschätzte konfirmatorische Faktoren- bzw. Strukturgleichungsmodelle

potentieller Response Sets (insbesondere Akquieszenz und *Straight Lining*) keine invers formulierten Items enthält.

Zunächst wird über die Schätzung einer erneuten PCA die Stabilität der einfaktoriellen Struktur von K1 geprüft, wenn lediglich die sechs relevanten Items berücksichtigt werden (Lösung K1(6i) in Tabelle 2). Diese kann bestätigt werden, bei einem erwartungsgemäß stark erhöhten Eigenwertanteil von 40,9 %. Abgesehen von Item $i14_h$ (das einzige Item, das die MF *Fairness* repräsentiert) weisen alle Items Faktorladungen auf, die den in Abschnitt 4.4 festgelegten Schwellenwert $|\lambda_S| = 0,5$ deutlich überschreiten. Verzichtet man in einem weiteren Schritt auf Item $i14_h$ (Lösung K1(5i) in Tabelle 2), erhöht sich der Eigenwertanteil auf respektable 46,4 % bei konstanten Werten der verbleibenden Faktorladungen. Weiterhin darf die Reliabilität der Skala mit $\omega = 0,708$ als zufriedenstellend bewertet werden ($se_\omega = 0,032$; $CI95u = 0,645$; $CI95o = 0,770$) und die Anzahl von fünf Items entspricht exakt dem in Abschnitt 4.1 definierten Idealwert pro Skala. Vor diesem Hintergrund erscheint es aus empirischer Sicht als gerechtfertigt, die Items $i2_h$ (wird in der Folge zu $h1$ umbenannt), $i5_h$ ($h2$), $i6_h$ ($h3$), $i13_h$ ($h4$) und $i19_h$ ($h5$) als Skala zur Messung handlungsorientierter Moralvorstellungen (HMor) zu definieren.

Im Rahmen der Validierungsstudie (*Sample 2*; S2) kann die Performance der Skala überzeugend bestätigt werden: Der Eigenwertanteil liegt hier sogar bei 50,6 % und die Werte der Faktorladungen überschreiten mehrheitlich jene aus der Entwicklungsstichprobe (*Sample 1*; S1). Einzig das Item $i13_h$ erweist sich als kritisch, da es mit 0,44 den Schwellenwert der Faktorladungen moderat unterschreitet. Die interne Konsistenz der Skala entspricht mit $\omega = 0,766$ ($se_\omega = 0,020$; $CI95u = 0,723$; $CI95o = 0,799$) jener aus der Entwicklungsstichprobe.

Wendet man sich der Entwicklung der seinsorientierten Moralitätsskala (SMor) zu, so offenbaren die Ergebnisse einer ersten PCA unter Berücksichtigung aller 18 konstruierten Items (siehe Tabelle 9) eine vierdimensionale Datenstruktur (Tabelle 3). Auf der ersten Komponente laden insgesamt sieben Items ($i2_s$, $i4_s$, $i5_s$, $i6_s$, $i9_s$, $i16_s$, $i18_s$), der Eigenwertanteil beträgt 16,3 %. Inhaltlich bilden vier Items ($i4_s$, $i5_s$, $i9_s$, $i16_s$) die MF *Sanctity* ab, zwei Items ($i2_s$, $i18_s$) *Care* und ein Item ($i6_s$) *Fairness*. Die zweite extrahierte Komponente umfasst vier Items ($i12_s$, $i13_s$, $i15_s$, $i17_s$; Eigenwertanteil der Komponente: 11,6 %), wovon die drei letzten Items dem Delinquenzfaktor angehören, der zur explorativen Prüfung der diskriminanten Validität eingeführt wurde und auch das vierte Item ($i12_s$; Nichteinhaltung der Mülltrennung) repräsentiert die Akzeptanz eines Gesetzesverstößes, nämlich gegen das Kreislaufwirtschaftsgesetz. Die höheren Faktorladungen der Items $i13_s$, $i15_s$ und $i17_s$ im Vergleich zu Item $i12_s$ sind mutmaßlich auch darauf zurückzuführen, dass es sich bei diesen um verkehrsbezogene Delinquenzhandlungen handelt. Zusammenfassend lässt sich konstatieren, dass K2 erwartungsgemäß kein geeigneter Kandidat ist, um seinsorientierte Moralvorstellungen abzubilden. Vielmehr konnte im Sinne der diskriminanten Validität nachgewiesen werden, dass die anderen extrahierten Komponenten (insbesondere K1) andere latente Konstrukte repräsentieren als *Deviant Morality Scales* (DMS). Auf den beiden letzten extrahierten Komponenten laden jeweils nur zwei Items (K3: $i7_s$, $i11_s$; Eigenwertanteil der Komponente: 9,7 % || K4: $i8_s$, $i10_s$; Eigenwertanteil der Komponente: 8,6 %). Während sich die Items von K3 auf

Item	S1			S2		
	K1	K2	K3	K1(6i)	K1(5i)	K1(5i)
i1 _h						
i2 _h	0,674			0,661	0,700	0,778
i3 _h			0,597			
i4 _h		0,612				
i5 _h	0,712			0,737	0,756	0,739
i6 _h	0,718			0,764	0,786	0,783
i7 _h		0,523				
i8 _h			0,577			
i9 _h						
i10 _h			0,584			
i11 _h						
i12 _h			0,492			
i13 _h	0,444			0,594	0,570	0,440
i14 _h	0,422			0,449		
i15 _h		0,464				
i16 _h		0,541				
i17 _h						
i18 _h		0,559				
i19 _h	0,597			0,577	0,561	0,758
Eig.	2,735	2,142	1,725	2,451	2,319	2,532
Ant.	0,144	0,113	0,009	0,409	0,464	0,506

Anmerkungen: S1 ... *Sample 1* (Entwicklungsstichprobe); S2 ... *Sample 2* (Validierungsstichprobe); K ... Komponente; Eig. ... Eigenwert; Ant. ... Anteil der durch die Komponente gebundenen Gesamtvarianz aller Items; zur optimierten Darstellung der Itemzugehörigkeiten zu den extrahierten Komponenten wird auf den Ausweis der Faktorladungen mit $|\lambda| < 0,5$ verzichtet. Die grau hinterlegten Zeilen indizieren jene Items, die schließlich ausgewählt wurden.

Tabelle 2: PCA-basierte Faktorladungsmatrix der Items zur Erfassung allgemeiner handlungsorientierter Moralvorstellungen (HMor)

die MF *Authority* beziehen und konkret die Bewertung der moralischen Beschaffenheit des gegenwärtigen politischen Systems zum Gegenstand haben, scheint es sich bei K4 vermutlich um einen Methodenfaktor zu handeln. Die beiden Items weisen keinen offensichtlichen inhaltlichen Bezug auf ($i8_s$ repräsentiert die MF *Fairness* und $i10_s$ *Care*), allerdings gehören sie den insgesamt vier invers formulierten Items an. Somit scheidet schließlich auch K3 und K4 als ernsthafte Kandidaten aus, die seinsorientierte Dimension allgemeiner Moralvorstellungen zu repräsentieren.

Werden die sieben Items der verbleibenden K1 wiederholt einer PCA unterzogen, weisen sie erneut eine eindimensionale Struktur auf, nun mit klarerweise deutlich erhöhtem Eigenwertanteil von 40,1 % (siehe K1(7i) in Tabelle 3). Da Item $i2_s$ (FM: *Care*) allerdings den für die Faktorladungen festgelegten Schwellenwert ($|\lambda_S| = 0,5$) unterschreitet, wird es ausgeschlossen und die Analyse wiederholt (siehe K1(6i) in Tabelle 3). Die Befunde indizieren einen moderaten Anstieg des Eigenwertanteils auf 44,4 %, und auch die Faktorladungen aller sechs Items liegen über dem Schwellenwert. Da die Faktorladung von Item $i18_s$ jedoch merklich zu jenen der restlichen Items abfällt, wird eine finale PCA ohne das entsprechende Item geschätzt (siehe K1(5i) in Tabelle 3). Der Eigenwertanteil des nun fünf Items umfassenden Modells liegt mit 48,5 % annähernd bei jenem der HMor-Skala zuvor. Weiterhin weisen alle Items Faktorladungen in zufriedenstellender Höhe auf. Die Skalenreliabilität liegt mit $\omega = 0,735$ ($se_\omega = 0,024$; $CI95u = 0,688$; $CI95o = 0,782$) ebenfalls im Zielbereich und ist mit jener von HMor vergleichbar. Da mit jeweils fünf Items nun auch ein ausgewogenes Verhältnis der Skalenlänge mit HMor besteht, werden die Items $i4_s$ (wird in der Folge zu $s1$ umbenannt), $i5_s$ ($s2$), $i6_s$ ($s3$), $i9_s$ ($s4$) und $i16_s$ ($s5$) zu SMor fusioniert.

Im Rahmen der Validitätsprüfung lassen sich die Befunde mit *Sample 2* (S2) als Datengrundlage präzise replizieren. Die fünf Items weisen ebenfalls eine eindeutig ein-faktorielle Struktur mit annähernd gleich hohen Faktorladungen auf, die jeweils über dem Schwellenwert liegen. Folglich bestehen auch keine systematischen Unterschiede in der internen Konsistenz ($\omega = 0,728$; $se_\omega = 0,022$; $CI95u = 0,684$; $CI95o = 0,770$).

5.2 Validitätsprüfung und Abschätzung der Kontaminierung durch sozial erwünschtes Antwortverhalten

Im nächsten Analyseschritt werden die explorativ entwickelten Skalen HMor und SMor auf messtheoretischer Grundlage der *klassischen Testtheorie* mittels CFA auf ihre Passung an die jeweiligen empirischen Daten geprüft. Basierend auf den in Abschnitt 4.4 festgelegten Kriterien kann HMor in beiden Stichproben ein zufriedenstellender Fit attestiert werden (M1 in Tabelle 4). So liegt jeweils lediglich einer der Werte der drei Fit-Indizes marginal über/unter den definierten Schwellenwerten. Hinsichtlich der standardisierten Faktorladungen zeigt sich, dass das Item $h5$ (*...versuche ich nachhaltig zu sein*) in beiden Samples den Schwellenwert von $|\lambda_S| = 0,4$ unterschreitet und somit das *Problemitem* der Skala zu sein scheint. Für SMor (M2) divergiert die Modellgüte zwischen den beiden Stichproben deutlich. Während die Passung des konfirmatorischen Faktorenmodells an die Daten in *Sample 1* als gut bezeichnet werden kann, weist für *Sample 2* lediglich der SRMR einen angemessenen Fit auf. Dies ist offensichtlich

Item	S1				S2			
	K1	K2	K3	K4	K1(7i)	K1(6i)	K1(5i)	K1(5i)
i1 _s								
i2 _s	0,516				0,451			
i3 _s								
i4 _s	0,532				0,635	0,644	0,661	0,572
i5 _s	0,431				0,607	0,648	0,639	0,603
i6 _s	0,570				0,653	0,648	0,683	0,614
i7 _s			0,738					
i8 _s				0,747				
i9 _s	0,782				0,737	0,743	0,751	0,835
i10 _s				0,682				
i11 _s			0,783					
i12 _s		0,427						
i13 _s		0,691						
i14 _s								
i15 _s		0,763						
i16 _s	0,704				0,729	0,747	0,742	0,785
i17 _s		0,768						
i18 _s	0,547				0,575	0,560		
Eig.	2,930	2,079	1,744	1,542	2,806	2,665	2,426	2,379
Ant.	0,163	0,116	0,097	0,086	0,401	0,444	0,485	0,476

Anmerkungen: S1 ... *Sample 1* (Entwicklungsstichprobe); S2 ... *Sample 2* (Validierungsstichprobe); K ... Komponente; Eig. ... Eigenwert; Ant. ... Anteil der durch die Komponente gebundenen Gesamtvarianz aller Items; zur optimierten Darstellung der Itemzugehörigkeiten zu den extrahierten Komponenten wird auf den Ausweis der Faktorladungen mit $|\lambda| < 0,5$ verzichtet. Die grau hinterlegten Zeilen indizieren jene Items, die schließlich ausgewählt wurden.

Tabelle 3: Konzipierte Items zur Erfassung allgemeiner seinsorientierter Moralvorstellungen (SMor)

darauf zurückzuführen, dass die Faktorladungen drei Items s_1 (...*beschämt die Verschwendung von Lebensmitteln*) s_2 (*Schlechte Arbeitsbedingungen ... unzumutbar.*) und s_3 (...*Menschen mit sehr hohem Einkommen humanitäre Spendenaktionen unterstützen.*) jeweils moderat unter dem Schwellenwert liegen. Im Gegensatz dazu weisen die beiden verbleibenden Items (sehr) hohe Faktorladungen auf. Werden beide Skalen in einem Zwei-Faktorenmodell gemeinsam eingeführt (M3), zeigen sich zunächst im Vergleich zu M1 und M2 sehr ähnlich ausgeprägte Faktorladungen. Weiterhin indizieren, abgesehen vom CFI, die Fit-Indizes in beiden Stichproben einen angemessenen Fit. Die Faktorenkorrelationen (ϕ_{HS}) erweisen sich mit 0,606 (*Sample 1*) und 0,601 (*Sample 2*) als extrem stabil und liegen in einem theoretisch erwartbaren Wertebereich. Die empirisch ermittelte Stärke des Zusammenhangs zwischen HMor und SMor legt nahe, dass die beiden Skalen das zugrundeliegende latente Konstrukt ‘allgemeine Moralvorstellungen’ im Sinne einer *thick conception* handlungs- und seinsbezogen abbilden. Werden die Items der beiden Skalen allerdings zu einem gemeinsamen Faktor fusioniert (M4), reduziert sich der Modellfit deutlich und unterschreitet in beiden Stichproben die Akzeptanzgrenze, d.h. das Ein-Faktorenmodell reproduziert die Korrelationsstruktur der Items nicht zufriedenstellend. In Kombination mit den vorherigen Befunden zur Faktorenkorrelation manifestiert sich nun das messtheoretische Problem, dass beide Skalen zwar offensichtlich das gleiche (bzw. zumindest ein extrem ähnliches) latentes Konstrukt repräsentieren, dessen Komplexität allerdings die Spezifikation eines komplexeren Messmodells erfordert als jene des Ein-Faktorenmodells. Eine solche, wie z.B. das *Second Order CFA-Modell* (für Details siehe z.B. Reinecke (2014), ist mit zwei Skalen als latente Konstrukte auf der ersten Ebene aus Identifikationsgründen allerdings nicht möglich. Der praktische Umgang mit dieser Problematik wird in Abschnitt 6 diskutiert. Aus inhaltlicher Sicht wird abschließend mit Blick auf Tabelle 4 deutlich, dass der Gesamtfaktor in *Sample 1* klar von SMOR dominiert wird, während er in *Sample 2* einer deutlich ausgewogeneren Repräsentation beider Skalen entspricht, diesmal sogar mit leichtem Übergewicht zugunsten von HMor.

Nachdem die auf Grundlage von HMor und SMor spezifizierten CFA-Modelle auf ihren Fit an die empirische Datengrundlage geprüft wurden, erfolgt nun die Evaluation der Kriteriumsvalidität an typischen Alltagsdelikten der erwachsenen Allgemeinbevölkerung (S1: Schwarzfahren mit öffentlichen Verkehrsmitteln; S2: Trunkenheit am Steuer). Konkret wird die Prämisse der *Morality Theories* (auch wenn diese sich in ihrer Ausgestaltung unterscheiden) geprüft, dass ausgeprägte allgemeine Moralvorstellungen eine (alltags-)kriminalitätsreduzierende Wirkung besitzen. Das Hauptaugenmerk der Interpretation liegt auf den entsprechenden standardisierten Effektparametern. Beginnend mit dem Ein-Faktorenmodell für HMor (M5 in Tabelle 5) und SMor (M6) zeigt sich für beide Skalen in beiden Stichproben – und somit auch für beide Delikte – ein moderat bis mittelstark ausgeprägter Effekt in erwarteter negativer Richtung (S1: $\beta_H = -0,246$; $\beta_S = -0,157$ | S2: $\beta_H = -0,151$; $\beta_S = -0,135$). Dieser Befund spricht zunächst für die Kriteriumsvalidität der beiden Skalen. Weiterhin werden in M7 wiederum beide Faktoren gemeinsam eingeführt. Während in *Sample 1* SMor mehr oder weniger vollkommen von HMor absorbiert wird ($\beta_H = -0,246$; $\beta_S = -0,009$) und die Faktorenkorrelation im Vergleich mit dem zugrundeliegenden CFA-Modell (M3) unverändert verbleibt

Skala	Par.	M1		M2		M3		M4	
		S1	S2	S1	S2	S1	S2	S1	S2
HMor	λ_{h1}	0,620	0,736			0,589	0,733	0,413	0,681
	(se_{h1})	(0,047)	(0,035)			(0,057)	(0,348)	(0,112)	(0,036)
	λ_{h2}	0,736	0,665			0,698	0,637	-0,014	0,618
	(se_{h2})	(0,047)	(0,038)			(0,049)	(0,037)	(0,218)	(0,040)
	λ_{h3}	0,742	0,713			0,741	0,676	0,278	0,635
	(se_{h3})	(0,038)	(0,034)			(0,037)	(0,033)	(0,165)	(0,037)
	λ_{h4}	0,460	0,371			0,423	0,376	0,105	0,346
	(se_{h4})	(0,055)	(0,047)			(0,050)	(0,042)	(0,112)	(0,041)
λ_{h5}	0,424	0,707			0,410	0,748	0,141	0,720	
(se_{h5})	(0,061)	(0,039)			(0,064)	(0,038)	(0,099)	(0,037)	
SMor	λ_{s1}			0,543	0,428	0,500	0,593	0,439	0,587
	(se_{s1})			(0,054)	(0,060)	(0,062)	(0,053)	(0,065)	(0,041)
	λ_{s2}			0,545	0,448	0,505	0,469	0,538	0,468
	(se_{s2})			(0,058)	(0,055)	(0,056)	(0,046)	(0,064)	(0,046)
	λ_{s3}			0,599	0,450	0,598	0,439	0,553	0,414
	(se_{s3})			(0,042)	(0,052)	(0,041)	(0,051)	(0,051)	(0,051)
	λ_{s4}			0,684	0,923	0,672	0,891	0,607	0,806
	(se_{s4})			(0,044)	(0,039)	(0,042)	(0,033)	(0,058)	(0,029)
λ_{s5}			0,680	0,784	0,690	0,762	0,713	0,731	
(se_{s5})			(0,053)	(0,040)	(0,049)	(0,040)	(0,059)	(0,048)	
	ϕ_{HS}					0,606	0,601		
	se_{HS}					(0,065)	(0,054)		
Fit	χ^2	13,51	20,03	13,96	31,87	72,43	96,14	130,42	151,74
	(df)	(5)	(5)	(5)	(5)	(34)	(34)	(35)	(35)
	RMSEA	0,056	0,082	0,057	0,110	0,046	0,064	0,071	0,087
	CFI	0,883	0,910	0,910	0,854	0,790	0,809	0,480	0,641
	SRMR	0,046	0,048	0,038	0,077	0,052	0,080	0,162	0,141

Anmerkungen: M ... Modell; S1 ... *Sample 1* (Entwicklungsstichprobe); S2 ... *Sample 2* (Validierungsstichprobe); λ_* ... standardisierte Faktorladung des Items *; se ... Standardfehler. Die grau hinterlegten Felder indizieren standardisierte Faktorladungen $|\lambda| < 0,5$ bzw. Werte in den Fit Indizes, die die festgelegten Schwellenwerte nicht erreichen.

Tabelle 4: CFA-Schätzergebnisse der Messmodelle (M1–M4)

($\phi_{HS} = 0,614$), schwächen die beiden Faktoren in *Sample 2* im Vergleich zu M4 bzw. M5 ihre Einflüsse gegenseitig ab ($\beta_H = -0,127$; $\beta_S = -0,078$). Hiervon ist SMor ebenfalls stärker betroffen, und auch die Faktorenkorrelation reduziert sich im Vergleich zu M3 deutlich ($\phi_{HS} = 0,434$).²⁰ In M8 werden beide Faktoren analog zu M4 zu einem Gesamtfaktor fusioniert. Trotz des schlechten Fits lässt sich die Kriteriumsvalidität dieses Faktors empirisch belegen (S1: $\beta_G = -0,238$; S2: $\beta_G = -0,181$). Aufgrund der höheren Erklärungskraft von HMor weisen die Effektparameter ähnliche Werte zu jenen aus M5 auf.

Die Aufnahme der SAT-DMS in M9 stellt eine zentrale Modellerweiterung dar. Wie geplant soll damit einerseits eine diskriminante Validitätsprüfung von HMor und SMor erfolgen und somit deren Abgrenzung von den in Abschnitt 7 kritisierten *Deviant Morality Scales* evaluiert werden. Andererseits ist damit das Ziel verbunden, die Skalen eines komparativen Tests hinsichtlich Ihres Potentials zur Erklärung delinquenter Handlungen zu unterziehen. Wie sich zeigt, sind die beiden Skalen in *Sample 1* positiv mit der SAT-DMS korreliert (HMor: $\phi_{HD} = 0,291$; SMor: $\phi_{SD} = 0,401$). Prinzipiell sind, wie in Abschnitt 4.4 ausgeführt, positive Korrelationen zwar zu erwarten, diese sollten allerdings nicht substantiell (Schwellenwert: $|\phi| = 0,3$) sein. Vor diesem Hintergrund erweist sich der Zusammenhang zwischen SMor und der SAT-DMS als kritisch. Dies ist insofern auch erstaunlich, als dass aufgrund der Handlungsorientierung erwartet werden musste, dass HMor in höherem Maße mit der SAT-DMS korreliert als SMor. In *Sample 2* ist diskriminante Validität hingegen gegeben ($\phi_{HD} = 0,008$; $\phi_{SD} = -0,142$).²¹ Hinsichtlich der Erklärungskraft dominiert in *Sample 1* HMor neben SMor auch eindeutig über die SAT-DMS ($\beta_H = -0,251$; $\beta_S = -0,052$; $\beta_D = -0,160$). In *Sample 2* weist HMor demgegenüber die geringste Erklärungskraft auf, während SMor und die SAT-DMS auf vergleichbarem Niveau liegen ($\beta_H = -0,086$; $\beta_S = -0,131$; $\beta_D = -0,187$). Schließlich wird in M10 erneut der Gesamtfaktor anstelle der beiden Skalen modelliert. Während die Korrelation mit der SAT-DMS in *Sample 1* in einem kritischen Bereich liegt ($\phi_{GD} = 0,406$), kann in *Sample 2* kein linearer Zusammenhang nachgewiesen werden ($\phi_{GD} = -0,053$). Insgesamt liegt somit substantielle, jedoch nicht uneingeschränkte empirische Evidenz vor, dass sich ALLMOR mit den beiden Skalen HMor und SMor tatsächlich auch empirisch präzise von einer idealtypischen *Deviant Morality Scale* abgrenzt. Ein Vergleich der Effektparameter in M10 zeigt auf, dass der Gesamtfaktor (und somit HMor und SMor gemeinsam) in beiden Stichproben eine keinesfalls schlechtere Erklärungskraft besitzen als die SAT-DMS (S1: $\beta_G = -0,190$; $\beta_D = -0,135$ | S2: $\beta_G = -0,186$; $\beta_D = -0,179$). Weiterhin offenbaren die Befunde eine hohe Konstanz des Effektparameters des Gesamtfaktors über die beiden Stichproben und auch die hohe empirische Bedeutsamkeit von SMor in *Sample 2*. Zusammenfassend lässt sich somit festhalten, dass ALLMOR, obwohl es auf einem wesentlich allgemeineren und umfassenderen theoretischem Konzept basiert

²⁰Dieser Befund tritt nicht gänzlich unerwartet auf. Zumindest erscheint es plausibel, dass die handlungsorientierten Moralvorstellungen einen stärkeren Einfluss auf das (selbstberichtete) Handeln der Akteure hat als seinsorientierte Moralvorstellungen. Allerdings erscheint es unter theoretischen Gesichtspunkten verkürzt, lediglich auf erstere abzustellen (siehe die Argumente in Abschnitt 2).

²¹Die negative Korrelation zwischen SMor und der SAT-DMS ist bei einem Test auf dem üblichen 5 %-Niveau statistisch nicht signifikant.

als die auf die moralische Bewertung devianten bzw. delinquenten Handelns verengten *Deviant Morality Scales*, diesen hinsichtlich des Potentials zur Erklärung entsprechender Handlungen zumindest ebenbürtig ist.

Tabelle 5: SEM-Schätzergebnisse zur Prüfung der Kriteriumsvalidität (M5–M10)

Par.	M5		M6		M7		M8		M9		M10	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
β_H (seH)	-0,246 (0,047)	-0,151 (0,052)	-0,246 (0,070)	-0,127 (0,061)	-0,251 (0,070)	-0,086 (0,062)	-0,238 (0,046)	-0,181 (0,052)	-0,190 (0,053)	-0,186 (0,052)	-0,190 (0,053)	-0,186 (0,052)
β_S (seS)		-0,135 (0,049)	-0,009 (0,072)	-0,078 (0,058)	-0,052 (0,075)	-0,131 (0,060)						
β_G (seG)												
β_D (seD)												
ϕ_{HS} (seHS)			0,614 (0,04)	0,434 (0,049)	0,604 (0,045)	0,450 (0,049)						
ϕ_{HD} (seHD)					0,291 (0,055)	0,008 (0,062)						
ϕ_{SD} (seSD)					0,401 (0,521)	-0,142 (0,057)						
ϕ_{GD} (seGD)											0,406 (0,050)	-0,053 (0,061)

*** Fortsetzung auf nächster Seite ***

Tabelle 5: SEM-Schätzergebnisse zur Prüfung der Kriteriumsvalidität (M5–M10)

*** Fortsetzung ***

Par.	M5		M6		M7		M8		M9		M10	
	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2	S1	S2
χ^2	32,49	31,62	29,95	56,34	112,93	142,81	258,30	438,33	381,21	326,08	534,46	616,83
(df)	(9)	(9)	(9)	(9)	(42)	(42)	(44)	(44)	(130)	(130)	(133)	(133)
RMSEA	0,069	0,075	0,066	0,109	0,056	0,074	0,096	0,143	0,061	0,059	0,076	0,092
CFI	0,955	0,956	0,959	0,913	0,938	0,911	0,813	0,650	0,867	0,887	0,787	0,721
SRMR	0,039	0,040	0,034	0,067	0,045	0,065	0,066	0,092	0,053	0,058	0,062	0,072

Anmerkungen: M... Modell; S1... *Sample 1* (Entwicklungsstichprobe); S2... *Sample 2* (Validierungsstichprobe); β_* ... standardisierter Effektparameter von * auf C; ϕ_{*+} ... Korrelationsparameter zwischen * und +; se... Standardfehler. Die grau hinterlegten Felder indizieren Werte in den Fit Indizes, die die festgelegten Schwellenwerte nicht erreichen. Auf den Ausweis der Parameter der Messmodelle wird aus Platzgründen verzichtet. Diese können bei Interesse von den Autoren angefordert werden.

Die beiden finalen Modelle (Tabelle 6) dienen der Bewertung der Skalen hinsichtlich ihrer Anfälligkeit für sozial erwünschtes Antwortverhalten. Während in M11 die Stärke des linearen Zusammenhangs von der sozialen Erwünschtheitsskala mit HMor und SMor in *Sample 1* auf vergleichbar hohem Niveau liegt ($\phi_{HE} = 0,444$; $\phi_{SE} = 0,488$), übersteigt der Wert der entsprechenden Korrelation mit der SAT-DMS dieses noch deutlich ($\phi_{DE} = 0,582$). In *Sample 2* sind die Zusammenhänge deutlich abgeschwächt,²² wobei sich nun die Korrelation der sozialen Erwünschtheitsskala mit SMor deutlich nach unten absetzt ($\phi_{SE} = 0,176$; im Vergleich: $\phi_{HE} = 0,321$; $\phi_{SE} = 0,330$). Vergleichbare Befunde liegen auch für das Gesamtfaktormodell M12 vor (S1: $\phi_{GE} = 0,537$; $\phi_{DE} = 0,583$ | S2: $\phi_{GE} = 0,346$; $\phi_{DE} = 0,330$).

Obwohl alle drei Skalen offensichtlich in deutlichem Maße sozial erwünschtes Antwortverhalten evozieren, erweisen sich die interessierenden Effektparameter davon jedoch als weitgehend unbeeinflusst, d.h. sie bleiben nach statistischer Kontrolle durch die jeweilige soziale Erwünschtheitsskala im Vergleich zu den Baseline-Modellen M9 bzw. M10 in beiden Stichproben erstaunlich robust (M11: S1: $\beta_H = -0,275$; $\beta_S = -0,051$; $\beta_D = -0,186$ | S2: $\beta_H = -0,153$; $\beta_S = -0,062$; $\beta_D = -0,153$ || M12: S1: $\beta_G = -0,215$; $\beta_D = -0,158$ | S2: $\beta_G = -0,174$; $\beta_D = -0,148$). Wiederum performt ALLMOR keineswegs schlechter als die SAT-DMS. Vielmehr das Gegenteil ist der Fall: Tendenziell weisen die Skalen eine geringere Korrelation mit den sozialen Erwünschtheitsskalen auf als die SAT-DMS und auch die Effektparameter des Gesamtfaktors liegen nach Berücksichtigung der Tendenz zu sozial erwünschtem Antwortverhalten (zum Teil deutlich) über jenen der SAT-DMS.

6 Diskussion

Dem Beitrag liegt die Absicht zugrunde, mit ALLMOR ein kompaktes zweidimensionales (handlungs- und seinsorientiert) Instrument zur Messung allgemeiner Moralvorstellungen zu entwickeln, das auf einem starken theoretischen Fundament basiert und die Limitierungen der bisher standardmäßig in der einschlägigen empirischen Forschung zum Einsatz gebrachten *Deviant Morality Scales* überwindet. Wie sich zeigt, weisen die beiden auf jeweils fünf selektierten Items basierenden Skalen HMor und SMor in der empirischen Prüfung eine weitgehend zufriedenstellende Skalenqualität hinsichtlich der Kriterien Reliabilität (interne Konsistenz), Kriteriumsvalidität, diskriminante Validität und Kontaminierung durch sozial erwünschtes Antwortverhalten auf. Insbesondere konnte in der Validierungsstichprobe (*Sample 2*) die jeweils eindimensionale Faktorstruktur überzeugend repliziert werden und aus beiden Stichproben lassen sich sehr ähnliche Konklusionen ableiten. So konnte im konkurrierenden Vergleich ebenfalls gezeigt werden, dass die eingesetzte SAT-DMS keinesfalls besser performt als ALLMOR.

Dennoch müssen abschließend einige Aspekte kritisch reflektiert werden. (i) Empirische Grundlage der Skalenentwicklung und -validierung waren Stichproben aus deutschen Großstädten. Aufgrund der Unterrepräsentanz ländlicher Regionen, aber auch Ostdeutschlands, sind die Befunde nicht repräsentativ für die gesamtdeutsche Allge-

²²Dies ist vermutlich auch darauf zurückzuführen, dass in *Sample 2* eine andere soziale Erwünschtheitsskala verwendet wurde als in *Sample 1* (siehe dazu Abschnitt 4.3).

Par.	M11		M12	
	S1	S2	S1	S2
β_H	-0,275	-0,153		
(se _H)	(0,074)	(0,071)		
β_S	-0,051	-0,062		
(se _S)	(0,078)	(0,064)		
β_G			-0,215	-0,174
(se _G)			(0,062)	(0,068)
β_D	-0,186	-0,153	-0,158	-0,148
(se _D)	(0,071)	(0,068)	(0,070)	(0,068)
β_E	0,056	-0,167	0,059	-0,169
(se _E)	(0,082)	(0,076)	(0,044)	(0,077)
ϕ_{HS}	0,616	0,367		
(se _{HS})	(0,045)	(0,061)		
ϕ_{HD}	0,293	-0,009		
(se _{HD})	(0,056)	(0,072)		
ϕ_{SD}	0,418	-0,123		
(se _{SD})	(0,052)	(0,068)		
ϕ_{GD}			0,409	-0,029
(se _{GD})			(0,050)	(0,072)
ϕ_{HE}	0,444	0,321		
(se _{HE})	(0,056)	(0,070)		
ϕ_{SE}	0,488	0,176		
(se _{SE})	(0,053)	(0,070)		
ϕ_{GE}			0,537	0,346
(se _{GE})			(0,049)	(0,069)
ϕ_{DE}	0,582	0,330	0,583	0,330
(se _{DE})	(0,051)	(0,070)	(0,051)	(0,070)
χ^2	670,71	587,29	816,94	829,33
(df)	(243)	(341)	(247)	(345)
RMSEA	0,058	0,049	0,067	0,068
CFI	0,830	0,856	0,773	0,717
SRMR	0,058	0,066	0,062	0,69

Anmerkungen: M ... Modell; S1 ... *Sample 1* (Entwicklungsstichprobe); S2 ... *Sample 2* (Validierungsstichprobe); β_* ... standardisierter Effektparameter von * auf C; ϕ_{*+} ... Korrelationsparameter zwischen * und +; se ... Standardfehler. Die grau hinterlegten Felder indizieren Werte in den Fit Indizes, die die festgelegten Schwellenwerte nicht erreichen. Auf den Ausweis der Parameter der Messmodelle wird aus Platzgründen verzichtet. Diese können bei Interesse von den Autoren angefordert werden.

Tabelle 6: SEM-Schätzergebnisse zum Einfluss sozialer Erwünschtheit (M11–M12)

meinbevölkerung. (ii) Im Rahmen der PCA-basierten Itemselektion reduzierte sich in beiden Skalen die Anzahl der enthaltenen *Moral Foundations* zum Teil deutlich. Somit ist die *Moral Domain* bedauerlicherweise nicht in jener Gesamtheit erfasst, die die *Moral Foundations Theory* vorgibt. Allerdings sind mit *Care* und *Fairness* jene beiden MFs prominent vertreten, die Priorität besitzen. (iii) Die finalen Fassungen der beiden Skalen enthalten keine der invers formulierten Items. Eine einfache Identifikation von Response Sets wie z.B. Akquieszenz und *Straight Lining* ist somit nicht möglich. (iv) Nicht alle Items besitzen im Rahmen der konfirmatorischen Analysen sehr hohe Faktorladungen. Dies trifft insbesondere auf Item *h4* (... *achte ich darauf niemanden zu verletzen oder zu beleidigen*) zu. (v) Letztlich konnten die empirischen Analysen keinen substantiellen Beitrag zum praktischen Umgang mit den beiden Skalen leisten. Gegen die Bildung eines Gesamtfaktors spricht der Fit der entsprechenden CFA-Modelle. Allerdings weisen die beiden Skalen einen sehr hohen Zusammenhang auf, was wiederum dafür spricht, dass sie das gleiche latente Konstrukt (auf höherer Ebene) repräsentieren. Weiterhin besitzt der Gesamtfaktor ein erhebliches Potential zur Erklärung von deviantem bzw. delinquentem Handeln. Diesen Aspekt gilt es weiterführend zu bearbeiten.

Literatur

- AAPOR. (2016). *Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys* (9. Aufl.). AAPOR.
- Abend, G. (2012). What the science of morality doesn't say about morality. *Philosophy of the Social Sciences*, 43(2), 157–200.
- Allen, C., Smit, I. & Wallach, W. (2005). Artificial morality: Top-down, bottom-up, and hybrid approaches. *Ethics and information technology*, 7(3), 149–155.
- Allen, C., Varner, G. & Zinser, J. (2000). Prolegomena to any future artificial moral agent. *Journal of Experimental & Theoretical Artificial Intelligence*, 12(3), 251–261.
- Antonaccio, O. & Tittle, C. R. (2008). Morality, Self-control, and Crime. *Criminology*, 46(2), 479–510.
- Billiet, J. (2016). What does measurement mean in a survey context? In C. Wolf, D. Joye, S. T. W. & Y.-c. Fu (Hrsg.), *The SAGE Handbook of Survey Methodology* (S. 193–209). Thousand Oaks, Sage.
- Brauer, J. R. & Tittle, C. R. (2017). When Crime is not an Option: Inspecting the Moral Filtering of Criminal Action Alternatives. *Justice Quarterly*, 34(5), 818–846.
- Brown, T. A. (2006). *Confirmatory Factor Analysis for Applied Research*. New York, Guilford Press.
- Browne, M. W. (1982). Covariance structures. In D. M. Hawkins (Hrsg.), *Topics in Applied Multivariate Analysis* (S. 72–142). Cambridge, Cambridge University Press.
- Browne, M. W. (1984). Asymptotically distribution-free methods for the analysis of covariance structures. *British Journal of Mathematical & Statistical Psychology*, 37(1), 62–83.
- Cohen, L. E. & Felson, M. (1979). Social Change and Crime Rate Trends: A Routine Activity Approach. *American Sociological Review*, 44(4), 588–608.
- Cooper, N. (1966). The two concepts of morality. *Philosophy*, 41(155), 19–33.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Dillman, D. A., Smyth, J. D. & Christian, L. M. (2014). *Internet, Phone, Mail, and Mixed-Mode Surveys. The Tailored Design Method*. Hoboken, Wiley.
- Dinno, A. (2009). Implementing Horn's parallel analysis for principal component analysis and factor analysis. *The Stata Journal*, 9(2), 291–298.
- Dunn, T. J., Baguley, T. & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, 105(3), 399–412.
- Durkheim, E. (1992 [1893]). *Über soziale Arbeitsteilung. Studie über die Organisation höherer Gesellschaften*. Frankfurt, Suhrkamp.
- Eifler, S. (2009). *Kriminalität im Alltag. Eine handlungstheoretische Analyse von Gelegenheiten*. Wiesbaden, VS Verlag.
- Eifler, S. & Leitgöb, H. (2018). Handlungstheoretische Ansätze zur Erklärung von Kriminalität. Eine Darstellung aus der Perspektive der analytischen Soziologie. In D. Hermann & A. Pöge (Hrsg.), *Kriminalsoziologie* (S. 9–38). Baden-Baden, Nomos.

- Festinger, L. (1957). *A Theory of Cognitive Dissonance*. Stanford, Stanford University Press.
- Frankena, W. K. (1966). The concept of morality. *Journal of Philosophy*, 63(21), 688–696.
- Gert, B. & Gert, J. (2017). The definition of morality. In E. N. Zalta (Hrsg.), *Stanford Encyclopedia of Philosophy* (Fall 2017 edition). Stanford, Stanford University. <https://plato.stanford.edu/entries/morality-definition>
- Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S. P. & Ditto, P. H. (2013). Moral foundations theory: The pragmatic validity of moral pluralism. *Advances in Experimental Social Psychology*, 47, 55–130.
- Grasmick, H. G. & Bursik, R. J. (1990). Conscience, significant others, and rational choice: Extending the deterrence model. *Law & Society Review*, 24, 837–862.
- Grasmick, H. G. & Green, D. E. (1980). Legal punishment, social disapproval and internalization as inhibitors of illegal behavior. *Journal of Criminal Law & Criminology*, 71, 325–335.
- Haidt, J. (2003). The moral emotions. In R. J. Davidson, K. R. Scherer & H. H. Goldsmith (Hrsg.), *Handbook of Affective Sciences* (S. 852–870). Oxford, Oxford University Press.
- Haidt, J. (2012). *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York, Random House.
- Haidt, J. & Joseph, C. (2004). Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, 133(4), 55–66.
- Haksar, V. (1998). Moral agents. In E. Craig (Hrsg.), *Routledge Encyclopedia of Philosophy* (S. 499–504). London, Routledge.
- Hitlin, S. & Stephen, V. (Hrsg.). (2010). *Handbook of the Sociology of Morality*. New York, Springer.
- Hitlin, S. & Vaisey, S. (2013). The new sociology of morality. *Annual Review of Sociology*, 39, 51–68.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2), 179–185.
- Johnson, D. G. (2006). Computer systems: Moral entities but nor moral agents. *Ethics & Information Technology*, 8(4), 195–204.
- Kant, I. (1990 [1797]). *Die Metaphysik der Sitten*. Ditzingen, Reclam.
- Kant, I. (1961 [1787]). *Kritik der praktischen Vernunft*. Ditzingen, Reclam.
- Kelley, K. (2019). *The MBESS R Package* (Techn. Ber.). <http://CRAN.R-project.org/package=MBESS>
- Kemper, C. J., Beierlein, C., Bensch, D., Kovaleva, A. & Rammstedt, B. (2012). *Eine Kurzsкала zur Erfassung des Gamma-Faktors sozial erwünschten Antwortverhaltens: Die Kurzsкала Soziale Erwünschtheit-Gamma (KSE-G)* (GESIS-Working Papers Nr. 25). Mannheim.
- Kline, R. B. (2005). *Principles and Practice of Structural Equation Modeling* (2. Aufl.). New York, Guilford Press.
- Kohlberg, L. (1974). *Zur kognitiven Entwicklung des Kindes*. Suhrkamp.
- Kohlberg, L. (1996). *Die Psychologie der Moralentwicklung*. Frankfurt, Suhrkamp.

- Kroneberg, C., Heintze, I. & Mehlkop, G. (2010). The interplay of moral norms and instrumental incentives in crime causation. *Criminology*, 48(1), 259–294.
- Leitgöb, H. & Eifler, S. (Hrsg.). (2018). Erklärungsangebote für kriminelles Handeln aus der Perspektive des Rational-Choice-Paradigmas. Eine Bestandsaufnahme [Sonderheft der Monatsschrift für Kriminologie und Strafrechtsreform], 101 (3/4).
- Luco, A. (2014). The definition of morality: Threading the needle. *Social Theory & Practice*, 40(3), 361–387.
- Matsueda, R. L. (2013). Rational choice research in criminology: A multi-level framework. In R. Wittek, T. A. B. Snijders & V. Nee (Hrsg.), *The Handbook of Rational Choice Social Research* (S. 283–321). Stanford, Stanford University Press.
- Mayo, B. (1968). The moral agent. *Royal Institute of Philosophy Supplements*, 1, 47–63.
- McDonald, R. P. (1999). *Test Theory: A Unified Treatment*. Mahwah, Lawrence Erlbaum Associates.
- Messner, S. F. & Rosenfeld, R. (2012). *Crime and the American Dream* (5. Aufl.). Wadsworth, Cengage Learning.
- Musch, J., Brockhaus, R. & Arndt, B. (2002). Ein Inventar zur Erfassung von zwei Faktoren sozialer Erwünschtheit. *Diagnostica*, 48(3), 121–129.
- Nagin, D. S. & Pogarsky, G. (2001). Integrating celerity, impulsivity, and extralegal sanction threats into a model of general deterrence: Theory and evidence. *Criminology*, 39(4), 865–892.
- Opp, K.-D. (1999). Contending conceptions of the theory of rational choice. *Journal of Theoretical Politics*, 11(2), 171–202.
- Opp, K.-D. (2013). Norms and rationality. Is moral behavior a form of rational choice? *Theory & Decision*, 74(3), 383–409.
- Padilla, M. A. & Divers, J. (2013). Coefficient Omega bootstrap confidence intervals: Nonnormal distributions. *Educational & Psychological Measurement*, 73(6), 956–972.
- Paternoster, R. & Simpson, S. (1996). Sanction Threats and Appeals to Morality: Testing a Rational Choice Model of Corporate Crime. *Law & Society Review*, 30(3), 549–583.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality & Social Psychology*, 46(3), 598–609.
- Paulhus, D. L. (1986). Self-deception and impression management in test responses. In A. Angleitner & J. S. Wiggins (Hrsg.), *Personality Assessment via Questionnaire* (S. 143–165). New York, Springer.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Hrsg.), *Measures of Personality and Social Psychological Attitudes* (S. 17–59). New York, Academic Press.
- Paulhus, D. L. (2002). Social desirable responding. The evolution of a construct. In H. I. Braun, D. N. Jackson & D. E. Wiley (Hrsg.), *The Role of Constructs in Psychological and Educational Measurement* (S. 49–69). Mahwah, Erlbaum.
- Pauwels, L. J. R., Svensson, R. & Hirtenlehner, H. (2018). Testing situational action theory: A narrative review of studies published between 2006 and 2015. *European Journal of Criminology*, 15(1), 32–55.

- Powell, C. (2010). Four concepts of morality. Differing epistemic strategies in the classical tradition. In S. Hitlin & S. Vaisey (Hrsg.), *Handbook of the Sociology of Morality* (S. 35–56). New York, Springer.
- Raykov, T. (1998). Coefficient alpha and composite reliability with interrelated nonhomogeneous items. *Applied Psychological Measurement*, *22*(4), 375–385.
- Reinecke, J. (2014). *Strukturgleichungsmodelle in den Sozialwissenschaften*. München, De Gruyter Oldenbourg.
- Saris, W. E. & Gallhofer, I. N. (2007). *Design, Evaluation, and Analysis of Questionnaires for Survey Research*. Hoboken, Wiley & Sons.
- Savage, L. J. (1954). *The Foundations of Statistics*. New York, Wiley.
- Schulz, S. & Kroneberg, C. (2018). Die situative Verursachung kriminellen Handelns – Zum Anwendungspotenzial des Modells der Frame-Selektion in der Kriminologie. *Monatsschrift für Kriminologie & Strafrechtsreform*, *101*(3-4), 251–271.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach’s alpha. *Psychometrika*, *74*(1), 107–120.
- Tangney, J. P., Stuewig, J. & Mashek, D. J. (2007). Moral emotions and moral behavior. *Annual Review of Psychology*, *58*(1), 345–372.
- Trasler, G. (1993). Routine Activity and Rational Choice. In R. V. Clarke & M. Felson (Hrsg.). New Brunswick, Transaction Publishers.
- Trizano-Hermosilla, I. & Alvarado, J. M. (2016). Best Alternatives to Cronbach’s Alpha Reliability in Realistic Conditions: Congeneric and Asymmetrical Measurements. *Frontiers in Psychology*, *7*, 769.
- Turiel, E. (1983). *The Development of Social Knowledge: Morality and Convention*. Cambridge, Cambridge University Press.
- Whiteley, C. H. (1960). On defining “moral”. *Analysis*, *20*(6), 141–144.
- Wikström, P.-O. (2006). Individuals, settings and acts of crime: Situational mechanisms and the explanation of crime. In W. Per-Olof & R. J. Sampson (Hrsg.), *The Explanation of Crime: Context, Mechanisms and Development* (S. 61–107). Cambridge, Cambridge University Press.
- Wikström, P.-O. H. (2010). Explaining crime as moral actions. In S. Hitlin & S. Vaisey (Hrsg.), *Handbook of the Sociology of Morality* (S. 211–239). New York, Springer New York.
- Wikström, P.-O. H., Oberwittler, D., Treiber, K. & Hardie, B. (2012). *Breaking rules: The social and situational dynamics of young people’s urban crime*. Oxford, Oxford University Press.
- Williams, B. A. O. (1985). *Ethics and the Limits of Philosophy*. Cambridge, Harvard University Press.
- Wolff, H.-G. & Bacher, J. (2010). Hauptkomponentenanalyse und explorative Faktorenanalyse. In C. Wolf & H. Best (Hrsg.), *Handbuch der sozialwissenschaftlichen Datenanalyse* (S. 333–365). Wiesbaden, VS Verlag.
- Yang, Y. & Green, S. B. (2011). Coefficient alpha: A reliability coefficient for the 21st century? *Journal of Psychoeducational Assessment*, *29*(4), 377–392.

- Zhang, Z. & Yuan, K.-H. (2016). Robust Coefficients Alpha and Omega and Confidence Intervals With Outlying Observations and Missing Data. *Educational & Psychological Measurement*, *76*(3), 387–411.
- Zinbarg, R. E., Revelle, W., Yovel, I. & Li, W. (2005). Cronbach's alpha, Revelle's beta, McDonald's omega: Their relations with each and two alternative conceptualizations of reliability. *Psychometrika*, *70*(1), 123–133.

7 Anhang

Für wie falsch (im Sinne von unrecht) halten Sie es, wenn eine Person Ihres Alters die folgenden Dinge tut?

- d1 Mit dem Fahrrad eine rote Ampel überfahren
 - d2 Aus Spaß eine Straßenlaterne zerstören
 - d3 Graffiti an eine Hauswand sprühen
 - d4 In einem Geschäft eine CD stehlen
 - d5 In ein Gebäude einbrechen um etwas zu stehlen
 - d6 Cannabis rauchen
 - d7 Eine Waffe oder Gewalt anwenden, um einer anderen Person Geld oder Gegenstände abzunehmen
-

Tabelle 7: Items der für die erwachsene Allgemeinpopulation adaptierten 7-Item SAT-DMS

Item	Formulierung	MF	Level
Im Zuge meines Handelns ...			
i1 _h	stehen für mich persönliche Konsequenzen im Vordergrund. (-)	C	I
i2 _h	denke ich auch an zukünftige Generationen. h1	C	III
i3 _h	lasse ich mich hauptsächlich von meinen Gefühlen leiten.	Ego	
i4 _h	kümmert es mich nicht, was andere über mich denken könnten.	Ego	
i5 _h	denke ich an die Folgen meines Handelns für die Menschen in meinem Umfeld. h2	C	II
i6 _h	berücksichtige ich, welche Konsequenzen mein Handeln für die Gesellschaft haben könnte. h3	C	III
i7 _h	ist es mir gleichgültig, ob ich gegen Gesetze verstoße. (-)	A	I
i8 _h	folge ich religiösen Prinzipien.	L	III
i9 _h	lasse ich mich nicht von meinem Gewissen leiten.	Ego	
i10 _h	versuche ich mich an dem zu orientieren, was meine Eltern mir beigebracht haben.	L	II
i11 _h	bin ich nicht nur auf meinen eigenen Vorteil bedacht.	L	I
i12 _h	denke ich auch daran, was andere von mir erwarten.	L	II
i13 _h	achte ich darauf, niemanden zu verletzen oder zu beleidigen. h4	C	II
i14 _h	versuche ich auch immer wieder mal uneigennützig zu sein.	F	II
i15 _h	gehe ich über Leichen, um meine Ziele zu erreichen. (-)	F	II
i16 _h	achte ich nicht darauf, auch andere zufriedenzustellen. (-)	F	I
i17 _h	versuche ich begangene Fehler wieder gut zu machen.	F	II
i18 _h	folge ich ausnahmslos meinen eigenen Prinzipien	Ego	
i19 _h	versuche ich nachhaltig zu sein. h5	C	III

Anmerkungen: MF ... *Moral foundations*; C ... *Care*; A ... *Authority*; L ... *Loyalty*; F ... *Fairness*; Ego ... *Egozentrismus*; I ... *präkonventionell*; II ... *konventionell*; III ... *postkonventionell*; grau hinterlegte Zeilen indizieren die Items der finalen HMor-Skala

Tabelle 8: Konzipierte Items zur Erfassung handlungsorientierter allgemeiner Moralvorstellungen (HMor)

Item	Formulierung	MF
i1 _s	Ich bin bestürzt über den hohen Konsum von Luxusartikeln.	S
i2 _s	Menschen sollten sich zur Organspende bereiterklären.	C
i3 _s	In Ausnahmefällen haben ich Verständnis für die Missachtung von Menschenrechten durch den Staat. (-)	A
i4 _s	Mich beschämt die Verschwendung von Lebensmitteln. s1	S
i5 _s	Schlechte Arbeitsbedingungen halte ich für unzumutbar. s2	S
i6 _s	Ich erwarte, dass Menschen mit sehr hohem Einkommen humanitäre Spendenaktionen unterstützen. s3	F
i7 _s	Ich bin erschüttert über die Korruption in unserem Land.	A
i8 _s	Mir ist es nicht wert, für fair gehandelte Waren mehr zu bezahlen. (-)	F
i9 _s	Ich bin bestürzt über das Leid von Kriegsflüchtlingen. s4	S
i10 _s	Es ist nicht notwendig, älteren Personen in öffentlichen Verkehrsmitteln den eigenen Platz anzubieten. (-)	C
i11 _s	Mich entsetzt der verschwenderische Umgang mit Steuergeldern.	A
i12 _s	Es ist schon in Ordnung, wenn man es mit der Mülltrennung nicht so genau nimmt. (-)	A
i13 _s	In Zeitnot halte ich es für gerechtfertigt, in öffentlichen Verkehrsmitteln schwarz zu fahren	Del.
i14 _s	Die Unterstützung von Tierschutzorganisationen halte ich für eine gute Sache.	C
i15 _s	Ich sehe es nicht so eng, als Fußgänger die Straße bei rot zu überqueren.	Del.
i16 _s	Das globale Ausmaß an Kinderarbeit schockiert mich. s5	S
i17 _s	Ich halte es für gerechtfertigt, wenn jemand, der in Eile ist, zu schnell mit dem Auto fährt.	Del.
i18 _s	Mich berührt es, wenn Leute Zivilcourage zeigen.	C

Anmerkungen: MF... *Moral foundations*; S... *Sanctity*; C... *Care*; A... *Authority*; F... *Fairness*; Del. ... Delinquenz; grau hinterlegte Zeilen indizieren die Items der finalen SMor-Skala

Tabelle 9: Konzipierte Items zur Erfassung seinsorientierter allgemeiner Moralvorstellungen (SMor)

Bisher erschienene Eichstätter Beiträge zur Soziologie

- Nr. 1: Knut Petzold, Thomas Brunner & Carlos Watzka: Determinanten der Zustimmung zur Frauenquote. Ergebnisse einer Bevölkerungsumfrage (12/2013)
- Nr. 2: Joost van Loon & Laura Unsöld: The Work and the Net: a Critical Reflection on Facebook-Research Methods and Optical Mediation (02/2014)
- Nr. 3: Stefanie Eifler: Projekt „Zusammenleben in der Stadt“ - Methodendokumentation (03/2014)
- Nr. 4: Basil Wiese: Intercultural Atmospheres – The Affective Quality of Gift Situations (09/2014)
- Nr. 5: Florian Mayr: Gibt es Krisen, und wenn ja wie viele? – Theoretisch-konzeptionelle Überlegungen zu einer Soziologie der Krise (12/2014)
- Nr. 6: Ramona Kay: Delinquente Opfer und viktimisierte Täter? Eine Analyse des Offending-Victimization-Overlap mit Hilfe der International Self-Report Delinquency Study (ISRD-2) (Working Paper) (05/2015)
- Nr. 7: Stefanie Eifler, Natalja Menold & Sara Pinkas: Metrische Eigenschaften des Semantischen Differentials zur Messung von Selbst- und Fremdbeschreibungen (11/2015)
- Nr. 8: Rémy Bocquillon & Joost van Loon: Soundlapse: The Will to Know as a Matter of Concern (05/2016)
- Nr. 9: Joost van Loon: Geworfenheit und Intensität: Ein Ansatz zum monistischen Denken mit Heidegger (11/2016)
- Nr. 10: Franziska Hodek: „Das is‘ Matip, mann!“ – Empirische Ideen zu Effekten von *Accounting* im Profifußball (02/2017)
- Nr. 11: Stefanie Eifler, Ramona Kay & Sara Pinkas: Projekt „Zusammenleben in der Stadt“ – Eine Sozialraumanalyse aus der Perspektive der Theorie der sozialen Desorganisation (05/2017)
- Nr. 12: Simon J. Charlesworth: Merlau-Ponty: Devaluation and Competence (11/2017)
- Nr. 13: Edda Mack: *fremd* vor dem Anderen: Ein Versuch der Herauslösung des Anderen aus dem Fremden (05/2018)
- Nr. 14: Hannah Bucher: Soziale Erwünschtheit im Faktoriellen Survey. Eine empirische Analyse der Präsentationsform von Vignetten (11/2018)
- Nr. 15: Moritz Wischert-Zielke: Schwarzer Block in Hamburg: Busfahrgäste in Angst. Eine Online-Teilnehmer*innenvideo-Analyse (05/2019)
- Nr. 16: Miryam Verbeek-Teres: Visuelle Stimuli in der Umfrageforschung. Ein Überblick über die Anwendung von Fotografien in Befragungen (11/2019)
- Nr. 17: Heinz Leitgöb, Stefanie Eifler & Julia Weymeirsch: Ein Instrument zur Erfassung allgemeiner Moralvorstellungen (ALLMOR) (05/2020)
- Nr. 18: Hannah Schwarzbach & Stefanie Eifler: Einflüsse der Präsentationsform eines faktoriellen Surveys zur Erfassung von Sicherheit im öffentlichen Raum (11/2020)