

Entstehungsbedingungen sozial erwünschten Antwortverhaltens

Eine experimentelle Onlinestudie zum Einfluss des Wording und des Kontexts bei unangenehmen Fragen

Von Ivar Krumpal und Anatol-Fiete Näher¹

Zusammenfassung: Umfragedaten können durch zahlreiche Fehlerquellen systematisch verzerrt sein. Eine bedeutsame Fehlerquelle sind systematische Antwortverzerrungen aufgrund sozialer Erwünschtheit. Der vorliegende Beitrag präsentiert Erklärungen sozial erwünschten Antwortverhaltens bei unangenehmen Fragen und diskutiert Methoden um sozial erwünschtes Antwortverhalten zu reduzieren. In der Literatur zur empirischen Sozialforschung findet sich häufig die These, dass *entschuldigende Frageformulierungen* (sog. „Wording-Effekte“) und die Wahl eines *permissiven Fragekontexts* (sog. „Kontext-Effekte“) Antwortverzerrungen aufgrund sozialer Erwünschtheit reduzieren würden. In Anknüpfung an die *Theorie kognitiver Dissonanz* werden die vermuteten Wording- und Kontexteffekte auf die Bereitschaft, sozial unerwünschtes Verhalten zu berichten, theoretisch rekonstruiert und durch ein *experimentelles Design* empirisch getestet (4 randomisierte Versuchsgruppen, Erhebung online mit Access-Panel, n=1176). Unsere Befunde zeigen, dass die häufig angemessene Wirkung sowohl von (1) entschuldigenden Formulierungen als auch (2) der Kontexteinbettung unangenehmer Fragen auf die Bereitschaft der Befragten, sozial unerwünschtes Verhalten zu berichten, wenig konsistent ausfällt. Es werden alternative Erklärungen für sozial erwünschtes Antwortverhalten diskutiert und empirisch überprüft. Hierbei zeigt sich, dass die *wahrgenommene soziale Norm* den stärksten und konsistentesten Effekt auf die Bereitschaft des Befragten ausübt, potentiell peinliches bzw. sozial unerwünschtes Verhalten zu berichten.

1 Einleitung

Wissenschaftlich fundierte Surveys stellen für die empirische Sozialforschung ein unverzichtbares Instrument zur Theorieprüfung dar. Allerdings können unberücksichtigte bzw. mangelhaft kontrollierte Fehler im datengenerierenden Prozess die Validität von Umfragedaten gefährden und zu Fehlschlüssen bei der Theorieprüfung führen. Neben Sampling- und Selbstselektionsproblemen resultieren insbesondere aus der sozialen Dimension eines Surveys und dem Befragungsprozess selbst zahlreiche Fehler und Verzerrungen (Tourangeau et al. 2000). Ein bedeutendes Problem stellen in diesem Zusammenhang fehlerhafte Antworten dar. So können beispielsweise Respondenten und Respondentinnen² die geforderten kognitiven Erinnerungsleistungen nicht erbringen oder mangelnde Bereitschaft zeigen, sozial unerwünschte Verhaltensweisen wahrheitsgemäß zu berichten. Zudem können Aspekte des Designs wie der Datenerhebungsmodus, der Einsatz spezieller Befragungstechniken oder Details von Frageformulierungen das Auftreten solcher Fehler und das Ausmaß der Verzerrungen in den Schätzungen moderieren. Bisherige Arbeiten der empirischen Sozial- und Umfrageforschung belegen im Hinblick auf die Gewinnung qualitativ hochwertiger Daten zahlreiche Probleme. So sind insbesondere Antworten auf Fragen heiklen Inhalts, sogenannte „sensitive“ oder „unangenehme Fragen“ (Van Koolwijk 1969), mit systematischen Messfehlern behaftet (Groves et

1 Die Autoren danken Thomas Voss und Karl-Dieter Opp für ihre Unterstützung. Der Beitrag ist im Rahmen des Forschungsprojekts „Asking Sensitive Questions: Possibilities and Limits of Randomized Response and Other Techniques in Different Survey Modes“ im Leipziger Teilprojekt VO 684 / 11 entstanden, das von der Deutschen Forschungsgemeinschaft im Rahmen des Schwerpunktprogramms 1292 „Survey Methodology“ gefördert wird.

2 Im Folgenden wird auf den Zusatz „/ -innen“ verzichtet. Dies dient ausschließlich der sprachlichen Vereinfachung.

al. 2004). Sensitive Fragen können private oder sozial unerwünschte Verhaltensweisen, illegale Handlungen oder unsoziale Einstellungen offenlegen und werden daher von den Befragten häufig als indiskret, unangenehm oder gar bedrohlich empfunden (Beyer / Krumpal 2010; Couatts / Jann 2011; Couatts et al. 2011; Jann et al. 2012; Krumpal 2011; Skarbek-Kozietulska et al. 2012; Tourangeau / Yan 2007). Befragte geben auf sensitive Fragen, wie z.B. Fragen zum Sexualverhalten, zur Einnahme von Drogen oder Psychopharmaka, zu kriminellen Aktivitäten sowie zu ausländerfeindlichen oder antisemitischen Einstellungen häufig falsche oder ausweichende Antworten. Systematische Messfehler in Antworten auf sensitive Fragen sind ausführlich durch Studien belegt, die selbstberichtetes Verhalten mit externen Validierungsdaten vergleichen: So zeigen Fu et al. (1995), dass im National Survey of Family Growth (NSFG) in den USA nur 52% derjenigen Abtreibungen berichtet wurden, die im gleichen Zeitraum tatsächlich in amerikanischen Abtreibungskliniken vorgenommen wurden (systematisches „Underreporting“ von negativ konnotiertem Verhalten). Die Studie von Belli et al. (2001) belegt verzerrte Antwortmuster auf Fragen zur Wahlbeteiligung. Der Vergleich von offiziellen Wahlstatistiken mit Antworten der American National Election Studies (ANES) ergab, dass im Zeitraum von 1964-1990 8-14% mehr Personen angaben, sich an einer US-Wahl beteiligt zu haben, als dies tatsächlich der Fall war (systematisches „Overreporting“ von positiv konnotiertem Verhalten).³

In der Lehrbuchliteratur zu Methoden der empirischen Sozialforschung bzw. zur *Survey Methodology* (Groves et al. 2004) werden häufig „Kunstregeln“ beschrieben, um sozial erwünschtes Antwortverhalten bei sensitiven Fragen zu reduzieren. Eine bekannte Empfehlung besteht darin, sensitive Verhaltensfragen so zu formulieren, dass das erfragte Verhalten entschuldigt wird. Dies wird in der Survey Methodology häufig als „loading“-Strategie bezeichnet (Sudman / Bradburn 1982). Nach Groves et al. (2004) fallen hierunter Formulierungen, die den Befragten dazu einladen, Fragen zu heiklen Themen wahrheitsgemäß zu beantworten. Konkrete Formulierungshinweise gibt bereits Barton (1958). Demnach ist bei der Formulierung sensitiver Fragen ein „everybody-does-it“-Ansatz zu wählen. Zudem können Frageformulierungen entwickelt werden, die den Eindruck erwecken, das erfragte Verhalten sei – erstens – von gesellschaftlich anerkannten Autoritäten erwünscht, – zweitens – sei aus guten oder nachvollziehbaren Gründen zustande gekommen, oder – drittens – werde vom Fragesteller schon vermutet (Sudman / Bradburn 1982). Vergleichbare Empfehlungen finden sich auch bei Fowler (1995). Skepsis gegenüber solchen Empfehlungen findet sich hingegen bei Diekmann (2003). In Hinblick auf die Wirksamkeit dieser Empfehlungen existiert nach Tourangeau und Yan (2007) eine Forschungslücke: „Surprisingly few studies have examined the validity of these recommendations to use ‘forgiving’ wording“ (Tourangeau / Yan 2007: 874).

Eine weitere prominente Empfehlung besteht darin, sensitive Fragen in einem inhaltlichen Kontext derart zu framen, dass deren Sensitivität nicht stärker betont wird, als dies erforderlich ist (Sudman / Bradburn 1982; Fowler 1995; Groves et al. 2004). Die Wirkungen dieser spezifischen Kontextuierungen sind bisher ebenfalls kaum erforscht (Tourangeau / Yan 2007). Gegenstand des vorliegenden Beitrags ist es deshalb, empirisch zu überprüfen, inwieweit die Empfehlungen sowohl zu entschuldigenden Formulierungen als auch zur Kontexteinbettung von sensitiven Fragen wirksam sind und den Respondenten in höherem Maße sozial unerwünschte Antworten entlocken. Zunächst werden nachfolgend auf Grundlage von Theorien kognitiver Dissonanz (vgl. Aronson 1999; Festinger 1957) und dem Inklusions- / Exklusionsmodell (Schwarz / Bless 1992, 2007) die vermuteten Wording- und Kontexteffekte auf die Bereitschaft, sozial unerwünschtes Verhalten zu berichten, theoretisch rekonstruiert. Anschließend werden die theoretischen Mechanismen in Hypothesen überführt und durch ein

3 Die Frage nach der Wahlbeteiligung ist insofern heikel, als die Verneinung dieser Frage sozial unerwünschtes politisches Desinteresse offenlegt.

experimentelles Design empirisch getestet. Abschließend werden die empirischen Befunde diskutiert und mit den wenigen bisher durchgeführten Experimenten (Catania et al. 1996; Holtgraves et al. 1997; Presser 1990; Tourangeau & Smith 1996) zu ähnlichen Fragestellungen verglichen.

2 Theorie

In der Survey Methodology werden „sensitive Fragen“ als eine der bedeutendsten Ursachen für sozial erwünschtes Antwortverhalten in Interviews angesehen (Becker 2006; Becker / Günther 2004; Diekmann 2003; Esser 1986; Krumpal 2010; Schnell et al. 1999; Stocké 2004; Tourangeau et al. 2000). Die Literatur zu diesem Thema ist entsprechend umfangreich. Dennoch konnte sich bisher keine einheitlich anerkannte Definition durchsetzen, die angibt, was unter „sensitiven Fragen“ genau zu verstehen sei (Sudman et al. 1996; Tourangeau et al. 2000). Nach Tourangeau / Yan (2007) lassen sich die zahlreichen Definitionsversuche zusammenfassen. Sensitive Fragen erfordern demnach Selbstauskünfte über Merkmalsausprägungen, die von den Befragten als sozial unerwünscht empfunden werden: „A question is sensitive when it asks for a socially undesirable answer, when it asks in effect that the respondent admits he or she has violated a social norm“ (Tourangeau / Yan 2007: 860). Wir definieren sensitive Verhaltensfragen⁴ als Fragen nach Verletzungen von sozialen Normen durch bestimmte Verhaltensweisen.⁵ So verlangt der US-amerikanische National Survey on Drug Use and Health (NSDUH) Selbstauskünfte zum Konsum illegaler Drogen und somit zu strafrechtlich relevantem Verhalten. Dies kann erhebliche Bedenken seitens der Befragten hinsichtlich der Offenlegung von illegalem, privatem, sowie sozial unerwünschtem Verhalten zur Folge haben. Diese Bedenken liegen insbesondere in der Furcht der Befragten vor möglichen negativen Konsequenzen und rechtlicher oder sozialer Sanktionierung begründet.

Auf Grundlage der bestehenden Forschungsliteratur lassen sich drei Mechanismen bestimmen, die sozial erwünschtem Antwortverhalten zugrunde liegen können: *Erstens* zeigen frühere Studien, dass sich die normativen Bewertungen bzw. Erwünschtheitswahrnehmungen eines erfragten Merkmals zwischen den Befragten unterscheiden können. Der wahrgenommene Sensitivitätsgrad, die „Trait Desirability“ (Stocké / Hunkler 2007), einer Frage variiert dementsprechend zwischen den Respondenten. So zeigen Wyner (1980), Nicotera (1996) und Huang et al. (1998), dass die Wahl sozial erwünschter Antwortoptionen bei sensitiven Fragen mit der Einschätzung der entsprechenden gesellschaftlichen Bewertung korreliert ist. Eine mögliche Ursache für Unterschiede in den Sensitivitätswahrnehmungen einer Frage besteht darin, dass sich Normen zwischen verschiedenen sozialen Gruppen unterscheiden können (Opp 2001 b; Johnson / van de Vijver 2002). *Zweitens* kann das Antwortverhalten von Befragten aufgrund von dauerhaft stabilen Persönlichkeitsmerkmalen wie dem Bedürfnis nach sozialer Anerkennung, das von der konkreten Befragungssituation unabhängig ist, variieren. Diese Merkmale bestehen einerseits in der Neigung von Respondenten, sozial erwünscht zu antworten, um soziale Anerkennung zu erzielen. Bezeichnet wird dies als das „Impression-Management“ eines Respondenten (Paulhus 1984). Unter „Self-Deception“ wird andererseits die Eigenschaft von Befragten verstanden, unbewusst sozial erwünschte Antworten zu geben, um ein positives Selbstbild aufrecht zu erhalten (Holtgraves 2004; Leak / Fish 1989; Nederhof 1985; Paulhus 1984, 2002; Randall / Fernandes 1991; Stocké / Hunkler 2007). Eine weit verbreitete Skala, die beide Merkmale gleichermaßen abbildet, ist die Marlowe / Crowne – Skala

4 Mit der Annahme systematischer Messfehler bei sensitiven Fragen ist die zusätzliche Annahme verbunden, dass das erfragte Merkmal potentiell fehlerfrei beobachtet werden kann (Tourangeau et al. 2000; Groves et al. 2004).

5 Für eine weiterführende Diskussion des Normbegriffs siehe insbesondere Opp (2001 a) sowie Voss (2001).

(Crowne / Marlowe 1960). In diesem Zusammenhang ist *drittens* die Diskretion während des Interviews bedeutend. Antwortverzerrungen aufgrund von „Impression-Management“ sind möglich, wenn die Angaben eines Respondenten durch andere Personen wahrgenommen werden können. Diese Bedingung entfällt in anonymen Befragungsmodi. Sozial erwünschte Antworten können in diesem Fall nur aufgrund der „Self-Deception“ von Respondenten zustande kommen. Allerdings ist es möglich, dass Respondenten der Gewährleistung der Anonymität in selbstadministrierten Interviews misstrauen, wenn technisch die Möglichkeit gegeben ist, die Identitäten der Respondenten mit ihren Angaben in Verbindung zu bringen. Dies ist beispielsweise im Rahmen von Online-Surveys der Fall (Coutts / Jann 2008: 12). Hier sollte zur Erklärung sozial erwünschten Antwortverhaltens zusätzlich das „Impression-Management“ berücksichtigt werden. Insgesamt kann jedoch den Antwortverzerrungen, die durch „Self-Deception“ motiviert sind, in selbstadministrierten Interviews ein relativ stärkeres Gewicht beigemessen werden als in Befragungen mit Interviewern.

Zusätzlich können Eigenschaften der Items das Befragtenverhalten beeinflussen. So weisen bereits Cantril / Wilks (1940) auf die Auswirkungen unterschiedlicher Frageformulierungen inhaltlich gleicher Fragen auf das Antwortverhalten der Respondenten hin. Diese werden in der Survey Methodology als „Wording-Effekte“ bezeichnet. Belegt sind diese Effekte beispielsweise bei Blair et al. (1977): Peinliche Fragen, die die Respondenten aufforderten, Angaben zu der Häufigkeit eines sozial unerwünschten Verhaltens zu machen, erzielten deutlich höhere Werte, wenn diese in Form von ausführlichen Formulierungen gestellt wurden. Die Annahme, dass sich der inhaltliche Kontext einer Frage auf das Antwortverhalten der Befragten auswirkt, ist in der Survey Methodology als „Kontexteffekt“ bekannt (Tourangeau et al. 2000). Demnach wurde in einer Studie von Strack et al. (1991) eine gezielt ambivalent gestellte Frage bezüglich eines finanziellen Beitrags zur eigenen Ausbildung von der studentischen Befragtengruppe in Abhängigkeit vom Inhalt der vorangestellten Frage entweder als zu zahlender oder als zu erhaltender Beitrag interpretiert.

Es liegen bisher nur wenige Erklärungsansätze vor, die es erlauben, sozial erwünschtes Antwortverhalten in Surveys systematisch als Spezialfälle einer allgemeinen Theorie zu rekonstruieren. Ein solcher allgemeiner theoretischer Ansatz, aus dem sich spezielle Hypothesen zu sozial erwünschtem Antwortverhalten ableiten lassen, ist die Theorie kognitiver Dissonanz nach Festinger (1957). Wir ziehen eine Variante der Dissonanztheorie heran, um spezielle Effekte von verschiedenen Formulierungen sensitiver Fragen theoretisch herzuleiten (vgl. Aronson et al. 1991; Aronson 1999). Zudem zeigen wir, dass Kontexteffekte durch das Inklusions- / Exklusionsmodell der Informationsverarbeitung in Interviews erklärbar sind (vgl. Schwarz / Bless 1992, 2007). In Verbindung mit der Dissonanztheorie lassen sich weitere Hypothesen ableiten, um die Empfehlungen hinsichtlich der Formulierung und des Kontexts sensitiver Fragen (Sudman / Bradburn 1982, Fowler 1995, Groves et al. 2004) in Bezug auf ihre Wirksamkeit und Konsistenz zu testen.

2.1 Theorie der kognitiven Dissonanz

Die Theorie der kognitiven Dissonanz beruht in ihrem Kern auf den folgenden Annahmen (vgl. Festinger 1957):

1. Individuen verfügen über bestimmte Wissens Elemente, sogenannte „Kognitionen“, die aus Informationen über sie selbst, ihrer Umwelt und ihr Verhalten bestehen.
2. Zwei Kognitionen χ_1 und γ_1 sind füreinander dann relevant, wenn aus χ_1 Folgerungen bezüglich γ_1 getroffen werden können und umgekehrt. Beispielsweise sind aus der Feststellung χ_1 (z.B. „Gestern habe ich Verhalten α begangen“) Folgerungen γ_1 (z.B. „Dieses Verhalten α ist mir peinlich.“) bezüglich dieser Feststellung ableitbar.

3. Relevante Kognitionen können inhaltlich zueinander in Widerspruch geraten. Festinger bezeichnet diesen Zustand als „kognitive Dissonanz“. Kognitionen, die miteinander im Einklang stehen, sind zueinander „konsonant“. Es ist davon auszugehen, dass Individuen in der Regel mehr als zwei salienten Kognitionen ausgesetzt sind. Die Stärke der Gesamtdissonanz lässt sich deshalb durch das Verhältnis aller salienten Kognitionen γ_n (wobei: $n = 2, 3, \dots, m$), die zu χ_1 dissonant sind, mit allen salienten Kognitionen χ_n , die zu χ_1 konsonant sind, beschreiben. Einen zusätzlichen Einfluss auf die Dissonanzstärke haben die individuell unterschiedlichen Bedeutungsgrade der beteiligten Kognitionen.
4. Individuen streben durch ein gezieltes kognitives Bemühen danach, Dissonanz zu reduzieren.

Das Motiv, die kognitive Dissonanz zu reduzieren, ergibt sich für ein Individuum in Abhängigkeit von der Dissonanzstärke. Nach Festinger (1957) besteht die Möglichkeit, Dissonanz zu reduzieren, grundsätzlich darin, eines oder mehrere Elemente einer dissonanten Beziehung zu ändern, um den Anteil dissonanter gegenüber dem Anteil konsonanter Kognitionen zu verringern. Dies kann anhand von unterschiedlichen Mechanismen geschehen. Welcher Mechanismus gewählt und welches Element des kognitiven Systems geändert wird, ist durch die Veränderungsresistenz einzelner Kognitionen bestimmt. Diese ist von der Menge der kognitiven Elemente abhängig, die mit der betreffenden Kognition konsonant sind (Festinger 1957; Frey / Gaska 1993). Mit der Anzahl der konsonanten Beziehungen steigt die Wahrscheinlichkeit, dass bei der Änderung einer dieser Kognitionen neue dissonante Beziehungen entstehen. Zum Zwecke einer Dissonanzreduktion ist eine Kognition somit umso schwerer zu verändern, je höher die Anzahl der Beziehungen ist, die mit ihr konsonant sind. Im konkreten Fall kann dies bedeuten, dass mit der Änderung eines kognitiven Elements für die betreffende Person sehr hohe psychische Kosten verbunden sind,⁶ oder die Person gegebenenfalls auf den Nutzen verzichten muss, der aus der Aufrechterhaltung einer bestimmten Kognition resultiert. Sofern angenommen wird, dass stets die Alternative gewählt wird, die mit den niedrigsten psychischen Kosten verbunden ist, wird diejenige Kognition zuerst verändert, deren Veränderungsresistenz am geringsten ist. Des Weiteren ist zu beachten, dass unter Umständen die Änderung einer Kognition faktisch nicht möglich ist (Festinger 1957). Es sind drei Mechanismen der Dissonanzreduktion zu unterscheiden: (1) *Addition von Kognitionen*: Es werden neue Elemente derart zu den bestehenden Kognitionen hinzugefügt, dass sich die Gesamtdissonanz der bestehenden Kognitionen verringert. Beispielsweise kann zu zwei dissonanten Kognitionen χ_1 („Ich rauche täglich eine Schachtel.“) und γ_1 („Rauchen ist ungesund.“) eine dritte Kognition χ_2 („Aber Rauchen entspannt mich.“) hinzugefügt werden. (2) *Subtraktion von Kognitionen*: Dieser Mechanismus besteht darin, die Anzahl der dissonanten Elemente zu verringern. Bei bestehender Einsicht einer Person in die Tatsache, wie ungesund es tatsächlich ist, zu rauchen (γ_1), kann sie es beispielsweise ignorieren und verdrängen, aber auch unterlassen (Frey / Gaska 1993), eine Schachtel Zigaretten pro Tag zu rauchen (χ_1). (3) *Substitution von Kognitionen*: Dissonante werden durch konsonante Kognitionen ersetzt. Um beim Beispiel des Rauchers zu bleiben: Dieser kann γ_1 („Rauchen ist ungesund.“) dadurch ersetzen, dass er lediglich Zeitungsartikel liest, die die Folgen des Rauchens herunterspielen (χ_2).

Ein wichtiges Problem der allgemeinen Version der Dissonanztheorie ist, dass die Anzahl aller füreinander relevanten Kognitionen, die in einer spezifischen Situation salient sind, nicht erschöpfend bestimmt werden kann (Frey / Gaska 1993). Deshalb lassen sich weder die Gesamtstärke der Dissonanz noch ein Mechanismus der Dissonanzreduktion eindeutig bestimmen. Präzise Vorhersagen sind somit nur unter der Angabe von Randbedingungen, anhand derer sich die beteiligten Kognitionen spezifizieren lassen, aus der Dissonanztheorie ableitbar (Frey / Gaska 1993). Eine bedeutende Weiterentwicklung der Dissonanztheorie nach Festinger

6 Festinger (1957) nennt an dieser Stelle die beiden Faktoren „Schmerz oder Verlust“.

(1957), die sich zur Erklärung von Wording-Effekten bei sensitiven Fragen anwenden lässt, besteht in dem von Aronson vertretenen „Hypocrisy“-Paradigma (vgl. Aronson et al. 1991; Aronson 1999). Diesem Ansatz liegt die folgende zusätzliche Annahme zugrunde: Kognitive Dissonanz entsteht genau dann, wenn die geteilte soziale Norm hinsichtlich eines Verhaltens (γ_1) mit der Erinnerung an das betreffende Verhalten (χ_1) in Widerspruch gerät (Stone et al. 1997, Fointiat 2008). Dies ist dann der Fall, wenn das Verhalten gegen die entsprechende soziale Norm verstößt (Aronson 1999; Fointiat 2008).⁷

Sensitive Verhaltensfragen wurden definiert als Fragen nach Verhaltensweisen der Respondenten, hinsichtlich derer eindeutige soziale Normen vorliegen. Sie erinnern Respondenten potentiell an sozial unerwünschte Verhaltensweisen. Dem „Hypocrisy“-Paradigma zufolge entsteht kognitive Dissonanz in Zusammenhang mit sensitiven Fragen somit genau dann, wenn die folgenden zwei Annahmen erfüllt sind. Die erste Annahme lautet: Der Respondent teilt eine Norm, die dem erfragten Verhalten widerspricht. Die zweite Annahme lautet: Das erfragte Verhalten liegt beim Respondenten tatsächlich vor. Wir nehmen an, dass für eine Teilmenge der Respondenten beide Annahmen erfüllt sind. Für diese Respondenten, kann die Leugnung eines erfragten Verhaltens bei sensitiven Fragen somit als Dissonanzreduktion betrachtet werden. Eine weitere Annahme ist, dass die *entschuldigende Formulierung* einer sensitiven Verhaltensfrage eine zusätzliche Kognition verfügbar macht, die zum erfragten Verhalten *konsonant* ist. Die induzierte kognitive Dissonanz ist deshalb bei entschuldigenden Formulierungen geringer als bei *direkten Formulierungen*. Entsprechend schwächer ist das Motiv zur Reduktion von Dissonanz. Sozial unerwünschtes Verhalten wird demnach in einem geringeren Ausmaß geleugnet. Es ergibt sich die folgende erste Hypothese:

Hypothese 1: Bei einer Formulierung, die ein sozial unerwünschtes Verhalten entschuldigt, steigt (verglichen mit einer direkten Formulierung) die Wahrscheinlichkeit, dass dieses Verhalten berichtet wird.

2.2 Das Inklusions- / Exklusionsmodell

Das Inklusions- / Exklusionsmodell nach Schwarz / Bless (1992, 2007) trägt aktuellen kognitionspsychologischen Befunden Rechnung (Greifeneder / Bless 2010). Hierbei werden die folgenden Annahmen getroffen:

1. Während der Beantwortung von Fragen bilden Individuen jeweils zwei „mentale Repräsentationen“ χ_1 und γ_1 (Schwarz / Bless 1992, 2007).⁸
2. χ_1 repräsentiert die Informationen, die zur Beantwortung einer Frage aus dem Gedächtnis abgerufen werden, γ_1 stellt einen Beurteilungsstandard dar, mit dem χ_1 verglichen wird. Das Schwarz / Bless-Modell (1992, 2007) wurde bisher nur explizit zur Erklärung von Kontexteffekten bei nicht-sensitiven Einstellungsfragen angewendet (Sudman et al. 1996, Tourangeau et al. 2000). Bei sensitiven Verhaltensfragen ist davon auszugehen, dass χ_1 die Erinnerung an das betreffende Verhalten und γ_1 das Wissen um die Existenz einer sozialen Norm hinsichtlich des Verhaltens repräsentiert.
3. Die beiden Repräsentationen χ_1 und γ_1 enthalten jeweils zwei Gruppen von Informationen, die entweder eine chronische oder eine zeitlich beschränkte Salienz aufweisen.
4. Sofern der Kontext einer Frage mit dieser thematisch verwandt ist, beeinflusst er die zeitlich beschränkt zugänglichen Informationen entweder in χ_1 oder γ_1 einer nachfolgenden Frage.

7 Die Annahme ist durch Aronson et al. (1991) und Stone et al. (1994) belegt. In beiden Studien wurde kognitive Dissonanz nur beobachtet, wenn sichergestellt war, dass die Respondenten eine eindeutige Norm hinsichtlich des Gebrauchs von Kondomen teilten, und wenn sie daran erinnert werden konnten, selbst gegen diese Norm verstoßen zu haben.

8 Die Begriffe „mentale Repräsentation“ (Schwarz / Bless 1992, 2007) und „Kognition“ (Festinger 1957) werden synonym verwendet.

Wenn der Fragekontext die Informationen beeinflusst, die zeitlich beschränkt in χ_1 *inkludiert* sind, resultieren sogenannte „Assimilationseffekte“ (Schwarz / Bless 1992, 2007). Die Antwortmuster einer nachgestellten Frage gleichen sich inhaltlich der zeitlich beschränkt salienten Information aus dem Fragekontext an. Tourangeau et al. (1989) zeigen, dass eine vorangestellte Frage nach der sozialen Verantwortung der amerikanischen Regierung in den Antworten der nachfolgenden Frage eine hohe Befürwortung öffentlicher Ausgaben zur Unterstützung wirtschaftlich benachteiligter Bevölkerungsgruppen nach sich zog. Für Befragte, die zuerst eine Frage zu wirtschaftlicher Eigenverantwortung beantworteten, galt dies hingegen nicht. Dagegen entstehen „Kontrasteffekte“, wenn der Fragekontext die zeitlich beschränkt salienten Informationen in γ_1 beeinflusst. Diese Informationen sind von χ_1 *exkludiert*. Die Antwortmuster sind demnach im Kontrast zum Inhalt der vorangestellten Frage verzerrt. In der Studie von Schwarz / Bless (1992) wurde zunächst ein Teil der Befragten dazu aufgefordert, Politiker zu benennen, die kurz zuvor in einen Skandal involviert gewesen waren. Anschließend sollten alle Teilnehmer die Glaubwürdigkeit eines Politikers bewerten, der an diesem Skandal nicht beteiligt gewesen war. Die Befragten bewerteten die Glaubwürdigkeit des Politikers signifikant höher, wenn ihnen zuerst die Frage nach den Politikern gestellt wurde, die für den Skandal verantwortlich gewesen waren.⁹

Die Vorhersagen des Schwarz / Bless-Modells (1992, 2007) lassen sich durch die Annahme eines Assimilationseffekts überprüfen: Nach Greifeneder / Bless (2010) ist dies möglich, indem man eine sensitive Verhaltensfrage in einem Kontext anordnet, dem eine bezüglich des erfragten Verhaltens liberale Einstellung zu entnehmen ist. Das „Hypocrisy“-Paradigma lässt sich mit dem Schwarz / Bless-Modell verknüpfen: Wir nehmen an, dass eine soziale Norm geteilt wird, die dem erfragten Verhalten widerspricht, und dass dieses Verhalten bei einer Teilmenge der Respondenten tatsächlich vorliegt. Des Weiteren nehmen wir an, dass der permissive Kontext einer sensitiven Verhaltensfrage eine zusätzliche Kognition verfügbar macht, die zum erfragten Verhalten *konsonant* ist. Die Stärke der induzierten kognitiven Dissonanz ist demzufolge geringer als im Falle eines neutralen Kontexts. Das Motiv zur Reduktion der Dissonanz ist entsprechend schwächer. Folglich ist zu erwarten, dass das sozial unerwünschte Verhalten im permissiven Kontext in einem geringeren Ausmaß geäußert wird:

Hypothese 2: Bei sensitiven Fragen mit einem permissiven Kontext steigt (verglichen mit einer direkten Formulierung) die Wahrscheinlichkeit, dass das sozial unerwünschte Verhalten berichtet wird.

Die Kunstregeln, den Kontext einer sensitiven Frage so zu gestalten, dass diese nicht in höherem Maße sensitiv auffällt als erforderlich (Sudman / Bradburn 1982; Fowler 1995; Groves et al. 2004), lassen sich ebenso unter der Annahme eines Assimilationseffekts auf ihre Wirksamkeit überprüfen. Es wird angenommen, dass der restriktive Kontext einer sensitiven Verhaltensfrage eine zusätzliche Kognition verfügbar macht, die zum erfragten Verhalten *dissonant* ist. Die Stärke der induzierten kognitiven Dissonanz ist demzufolge höher als im Falle eines neutralen Kontexts. Das Motiv zur Reduktion der Dissonanz ist entsprechend stärker. Folglich ist zu erwarten, dass das sozial unerwünschte Verhalten im restriktiven Kontext in einem höheren Ausmaß geäußert wird:

Hypothese 3: Bei sensitiven Fragen mit einem restriktiven Kontext sinkt (verglichen mit einer direkten Formulierung) die Wahrscheinlichkeit, dass das sozial unerwünschte Verhalten berichtet wird.

9 Ob Informationen in χ_1 inkludiert oder von χ_1 exkludiert werden, ist unter anderem auch dadurch bestimmt, wie leicht diese Informationen kognitiv zugänglich sind (Greifeneder / Bless 2010). Demnach führen leicht verfügbare Informationen zu Assimilationseffekten. Im Falle von schwer abrufbaren Informationen resultieren Kontrasteffekte. Greifeneder / Bless (2010) konnten diese Annahmen in zwei unabhängigen Experimenten belegen.

2.3 Bisherige Befunde

Es liegen bisher nur wenige Studien vor, die die Auswirkung von entschuldigenden Frageformulierungen bei sensitiven Fragen empirisch untersuchen. Hierbei zeigen sich inkonsistente Ergebnisse: In einer experimentellen Studie von Catania et al. (1996) wurden Respondenten telefonisch experimentell variierte Fragen zu ihrem Sexualverhalten gestellt.¹⁰ Während den Befragten in der Kontrollgruppe die Fragen in direkter Form gestellt wurden, waren die Fragen, die die Versuchsgruppe zu beantworten hatte, so formuliert, dass deren Intention weniger heikel erschien. Insgesamt zeigten sich einige positive Auswirkungen der variierten Frageformulierungen auf die Bereitschaft, sozial unerwünschte Antworten zu geben. Beispielsweise war die Zahl berichteter außerehelicher Affären oder des vorsätzlichen Nichtgebrauchs von Kondomen in der Versuchsgruppe signifikant höher als in der Kontrollgruppe. Eine weitere experimentelle Studie wurde von Holtgraves et al. (1997) durchgeführt.¹¹ Auch hier beantworteten Befragte in vier Experimenten unterschiedliche Frageversionen gleichen Inhalts. Den Fragen der Versuchsversionen ging ein Satz voraus, dem gute Gründe zu entnehmen waren, warum das entsprechende Verhalten begangen, bzw. unterlassen werden sollte. Holtgraves et al. (1997) zufolge zeigen sich keine Effekte für Fragen nach sozial erwünschtem Verhalten. Im Zusammenhang mit sozial unerwünschtem Verhalten belegen die Befunde nur teilweise die Effektivität einer entschuldigenden Frageformulierung. Belli et al. (2006) stellten per RDD („random digit dialing“) ausgewählten US-Bürgern in einer Telefonumfrage experimentell variierte Fragen zur Wahlbeteiligung. Frageformulierungen, denen gute Gründe für die Nichtbeteiligung an einer Wahl zu entnehmen waren, resultierten gegenüber der Kontrollgruppe in erwartungsgemäß niedrigeren Angaben zur Beteiligung an den letzten Kongresswahlen.

Kontexteffekte bei Umfragen sind mittels zahlreicher Studien empirisch belegt. Allerdings wurden die Auswirkungen des Kontextes explizit auf sensitive Fragen kaum experimentell überprüft. Die nachfolgenden Ergebnisse sind ebenfalls nicht konsistent und widersprechen teilweise dem Schwarz / Bless-Modell. In einer Studie von Presser (1990) wurden Respondenten telefonisch nach ihrer Beteiligung an den letzten Wahlen befragt.¹² Es wurden zwei Experimente mittels telefonischer Interviews durchgeführt. Im ersten Experiment wurde der Versuchsgruppe ein Fragekontext präsentiert, der aus einer Frage nach einer Information bestand, über die nur Wahlbesucher verfügen konnten. Es wurde angenommen, dass die Befragten in geringerem Maß dazu neigten, von einem Wahlbesuch zu berichten, wenn sie vorher angegeben hatten, dass sie über die Information nicht verfügten (Presser 1990). Der Versuchsgruppe im zweiten Experiment wurde ein Fragekontext vorgelegt, der aus einer Frage bestand, deren Beantwortung es den Befragten erlaubte, die Häufigkeit von Wahlbesuchen in der Vergangenheit anzugeben. Hiermit sollte den Befragten die Möglichkeit eingeräumt werden, die Nicht-Beteiligung an der letzten Wahl entschuldigend nur als einmaliges Ereignis darzustellen. In beiden Experimenten konnte Presser keine signifikanten Unterschiede in den Antwortmustern zwischen Kontroll- und Versuchsgruppen finden. In einem experimentellen Survey vari-

10 Die Auswahl der Untersuchungseinheiten erfolgte per RDD („random digit dialing“) unter US-Haushalten in verschiedenen US-Staaten.

11 Untersuchungseinheiten waren Psychologiestudenten einer US-amerikanischen Universität, die an den Experimenten teilnahmen, um Studienleistungen zu erbringen. Die Respondenten wurden schriftlich mittels selbstadministrierter Fragebögen interviewt.

12 Die Auswahl der Untersuchungseinheiten erfolgte zufällig per RDD aus US-Haushalten in einem US-Staat.

ierten Tourangeau / Smith (1996) den Kontext sensitiver Fragen wie folgt:¹³ Zwei Versuchsgruppen wurde eine Frage nach der Anzahl der bisherigen Sexualpartner vorgelegt. Dieser Frage wurde eine Einstellungsfrage vorangestellt, die die Respondenten aufforderte, ihre Bewertung jeweils einer restriktiven (Split 1) oder einer permissiven Aussage (Split 2) über promiskuitives Verhalten anzugeben. Im permissiven Kontext wurde eine höhere Anzahl berichteter Sexualpartner erwartet, als im restriktiven Kontext. Anders als erwartet zeigte sich der gegenteilige Effekt: Die restriktive Aussage wirkte sich positiv auf die Anzahl der berichteten Sexualpartner aus (Tourangeau / Smith 1996). Dieses Ergebnis ist nicht mit dem Schwarz / Bless-Modell konsistent, da bei gegebenem Versuchsaufbau Assimilationseffekte zu erwarten waren.

3 Methoden

Insgesamt werden vier Fragen experimentell variiert, von denen angenommen wird, dass sie als sensitiv einzustufen sind. Es handelt sich hierbei um dichotome Items mit den Ausprägungen „ja“ oder „nein“. In unserem experimentellen Survey sind dies die *abhängigen Variablen*:

Wahlbeteiligung: Ausgehend von Studien, bei denen Antworten zur Frage nach der Beteiligung an politischen Wahlen mit offiziellen Statistiken verglichen wurden (Traugott / Katosh 1979; Belli et al. 2001), werden für die direkte Frage zur Wahlbeteiligung, durch soziale Erwünschtheit verzerrte, hohe Wahlbeteiligungsraten erwartet (systematisches „Overreporting“ von sozial erwünschtem Verhalten). Für die drei anderen sensitiven Fragen werden dagegen Messfehler in die entgegengesetzte Richtung erwartet (systematisches „Underreporting“ von sozial unerwünschtem Verhalten).

Sexuelle Untreue in der Partnerschaft: Nach Catania et al. (1996) kann davon ausgegangen werden, dass Befragte dazu neigen, sexuelle Untreue in der Partnerschaft zu leugnen. Dies ist dann wahrscheinlich, wenn in Betracht gezogen wird, dass Treue in der Partnerschaft eine soziale Norm darstellt, die von vielen Personen geteilt wird. Zudem ist anzunehmen, dass die Frage von einigen Befragten als indiskret wahrgenommen wird und somit zusätzliches Underreporting verursacht. Es wird vermutet, dass einige Befragte potentiell daran Anstoß nehmen könnten und entsprechend sexuelle Untreue in der Partnerschaft nicht berichten werden.

Fahren unter Alkoholeinfluss: Der Frage nach dem Führen eines Kraftfahrzeugs unter starkem Alkoholeinfluss wird eine hohe Sensitivität unterstellt, da dieses Verhalten potentiell andere Personen gefährdet und strafbar ist.¹⁴ Es wird angenommen, dass Befragte dazu neigen, dieses Verhalten nicht zu berichten. Diese Annahme findet sich ebenfalls bei Coutts / Jann (2008).

Einnahme von Antidepressiva: Die Frage nach der ärztlichen Verordnung von Antidepressiva wird ebenso als sehr heikel angenommen.¹⁵ Die Verordnung von Psychopharmaka weist auf das Vorhandensein einer psychischen Erkrankung hin. Diese wird von den Betroffenen

-
- 13 Die Auswahl der Untersuchungseinheiten erfolgte durch Ziehung einer zweistufigen Zufallsstichprobe in einem US-Staat. Die Befragung erfolgte per ACASI („audio computer-assisted self-interviewing“), CASI („computer-assisted self-interviewing“) und CAPI („computer-assisted personal interviews“). Der Erhebungsmodus hatte in der Befragtengruppe mit restriktivem Fragekontext starke Effekte. Das höchste Ausmaß sozial erwünschten Antwortverhaltens wurde in CAPI beobachtet. Nur geringe Moduseffekte waren dagegen in der Gruppe mit permissivem Kontext zu verzeichnen.
 - 14 Eine Variante der Frage zu Fahren unter Alkoholeinfluss wurde beispielsweise im Teilmodul zur selbstberichteten Delinquenz im ALLBUS 1990 und 2000 gestellt.
 - 15 Fragebatterien zur Einnahme von Psychopharmaka lassen sich beispielsweise in dem regelmäßig in den USA durchgeführten National Survey on Drug Use and Health (NSDUH) finden.

häufig als stigmatisierend empfunden. Folglich ist davon auszugehen, dass Befragte dazu neigen, die Einnahme von Antidepressiva in Umfragen nicht zu berichten.

Abbildung 1: Direkte Formulierung

Wahlbeteiligung:

Waren Sie bei der letzten Bundestagswahl wählen?

Sexuelle Untreue in der Partnerschaft:

Sind Sie in Ihrer letzten / aktuellen festen Beziehung Ihrem Partner / Ihrer Partnerin sexuell untreu gewesen?

Fahren unter Alkoholeinfluss:

Haben Sie innerhalb der letzten zwei Jahre ein Kraftfahrzeug gesteuert, obwohl Sie die Promillegrenze überschritten hatten?

Einnahme von Antidepressiva:

Wurden Ihnen persönlich innerhalb der letzten zwei Jahre von einem Arzt Antidepressiva verschrieben?

Anmerkung: Die Items besitzen jeweils die Antwortoptionen „Ja“=1 und „Nein“=0

Im Folgenden werden die *Experimentalvariablen* beschrieben: Der Kontrollgruppe wurden die heiklen Fragen direkt gestellt (*Split 1: direkte Formulierung*). Um Hypothese 1 zu überprüfen, erhielt die erste Versuchsgruppe eine Version des Fragebogens, bei der jeweils die gleiche sensitive Frage eine Aussage enthält, die die erfragte Verhaltensweise entschuldigt (*Split 2: entschuldigende Formulierung*). Es werden plausible Gründe angegeben, warum diese Verhaltensweise vorliegen kann (vgl. Catania et al. 1996; Holtgraves et al. 1997). In der Version, die das erfragte Verhalten entschuldigt, werden mehr sozial unerwünschte „ja“-Antworten (bei Wahlbeteiligung „nein“-Antworten) erwartet als in der Version mit direkter Formulierung.

Abbildung 2: Entschuldigende Formulierung

Wahlbeteiligung:

Viele Leute haben nicht mehr die Zeit, um sich an politischen Wahlen zu beteiligen. Waren Sie bei der letzten Bundestagswahl wählen?

Sexuelle Untreue in der Partnerschaft:

Sozialwissenschaftler haben herausgefunden, dass immer mehr Menschen gelegentlich einen Seitensprung haben. Sind Sie in Ihrer letzten / aktuellen festen Beziehung Ihrem Partner / Ihrer Partnerin sexuell untreu gewesen?

Fahren unter Alkoholeinfluss:

Es kann schnell passieren, beim Genuss von Alkohol die für das Führen eines Kraftfahrzeugs zulässige Promillegrenze zu überschreiten. Haben Sie innerhalb der letzten zwei Jahre ein Kraftfahrzeug gesteuert, obwohl Sie die Promillegrenze überschritten hatten?

Einnahme von Antidepressiva:

Im heutigen Alltagsleben wird der Leistungsdruck immer größer. Viele wissenschaftliche Studien zeigen, dass deshalb depressive Erkrankungen zunehmen. Wurden Ihnen persönlich innerhalb der letzten zwei Jahre von einem Arzt Antidepressiva verschrieben?

Anmerkung: Die Items besitzen jeweils die Antwortoptionen „Ja“=1 und „Nein“=0

Die Operationalisierung der experimentellen Kondition zu Kontexteffekten erfolgt über die Bildung zweier weiterer Versuchsgruppen. Beiden Befragtengruppen wurden die unangenehmen Fragen in direkter Form gestellt. Diesen Fragen wurde ein zusätzliches Item vorgeschaltet, mittels dessen die Respondenten aufgefordert wurden, Aussagen (permissive versus restriktive) zu den nachfolgend erfragten Verhaltensweisen auf einer fünfstufigen Skala zu bewerten. Um Hypothese 2 zu testen, wurde der zweiten Versuchsgruppe eine Aussage vorgelegt, der eine liberale Einstellung in Hinblick auf das sensitive Merkmal zu entnehmen war (*Split 3: permissiver Kontext*). Im permissiven Kontext werden mehr sozial unerwünschte „ja“-Antworten (bei Wahlbeteiligung „nein“-Antworten) erwartet als in der Version mit direkter Formulierung.

Abbildung 3: Permissiver Kontext**Wahlbeteiligung:**

Stimmen Sie folgender Aussage zu? „Es bringt nichts, sich an politischen Wahlen zu beteiligen.“
Waren Sie bei der letzten Bundestagswahl wählen?

Sexuelle Untreue in der Partnerschaft:

Entspricht folgende Aussage Ihrer Meinung? „In einer Beziehung der Partnerin / dem Partner treu zu bleiben, ist unmöglich.“

Sind Sie in Ihrer letzten / aktuellen festen Beziehung Ihrem Partner / Ihrer Partnerin sexuell untreu gewesen?

Fahren unter Alkoholeinfluss:

Stimmen Sie dieser Aussage zu? „Auch oberhalb der Promillegrenze ist man noch in der Lage, Auto zu fahren.“

Haben Sie innerhalb der letzten zwei Jahre ein Kraftfahrzeug gesteuert, obwohl Sie die Promillegrenze überschritten hatten?

Einnahme von Antidepressiva:

Und teilen Sie folgende Meinung? „Die Einnahme von Antidepressiva kann in vielen Fällen die Lebensqualität erhöhen.“

Wurden Ihnen persönlich innerhalb der letzten zwei Jahre von einem Arzt Antidepressiva verschrieben?

Anmerkung: Die Items besitzen jeweils die Antwortoptionen „Ja“=1 und „Nein“=0. Die jeweils vorge-schalteten Kontextitems besitzen die Antwortoptionen 5=„stimme voll zu“, 4=„stimme zu“, 3=„teils / teils“, 2=„stimme nicht zu“, 1=„stimme überhaupt nicht zu“.

Um Hypothese 3 zu testen, wurde schließlich der dritten Versuchsgruppe eine Aussage vorgelegt, die eine bezüglich des nachfolgend erfragten Verhaltens restriktive Haltung zum Ausdruck brachte (*Split 4: restriktiver Kontext*). Im restriktiven Kontext werden weniger sozial unerwünschte „ja“-Antworten (bei Wahlbeteiligung „nein“-Antworten) erwartet als in der Version mit direkter Formulierung.

Abbildung 4: Restriktiver Kontext**Wahlbeteiligung:**

Stimmen Sie folgender Aussage zu? „Nur wer sich an politischen Wahlen beteiligt, zeigt auch demokratische Verantwortung.“

Waren Sie bei der letzten Bundestagswahl wählen?

Sexuelle Untreue in der Partnerschaft:

Entspricht die folgende Aussage Ihrer Meinung? „Wer sich dazu entscheidet, eine Beziehung zu führen, sollte seiner Partnerin / seinem Partner auch treu sein.“

Sind Sie in Ihrer letzten / aktuellen festen Beziehung Ihrem Partner / Ihrer Partnerin sexuell untreu gewesen?

Fahren unter Alkoholeinfluss:

Stimmen Sie folgender Aussage zu? „Leute, die betrunken Auto fahren, gefährden nicht nur sich selbst, sondern auch andere.“

Haben Sie innerhalb der letzten zwei Jahre ein Kraftfahrzeug gesteuert, obwohl Sie die Promillegrenze überschritten hatten?

Einnahme von Antidepressiva:

Und teilen Sie folgende Meinung? „Wer Antidepressiva einnimmt, kann persönliche Probleme nicht bewältigen.“

Wurden Ihnen persönlich innerhalb der letzten zwei Jahre von einem Arzt Antidepressiva verschrieben?

Anmerkung: Die Items besitzen jeweils die Antwortoptionen „Ja“=1 und „Nein“=0. Die jeweils vorge-schalteten Kontextitems besitzen die Antwortoptionen 5=„stimme voll zu“, 4=„stimme zu“, 3=„teils / teils“, 2=„stimme nicht zu“, 1=„stimme überhaupt nicht zu“.

Als wichtige Einflussgrößen auf das Antwortverhalten sind außerdem in allen vier Versionen enthalten: Die beiden Teildimensionen des Bedürfnisses nach sozialer Anerkennung, „Impression-Management“ und „Self-Deception“, wurden auf Grundlage einer deutschsprachigen Kurzversion der Marlowe / Crowne-Skala (Crowne / Marlowe 1960; Stocké 2007) operationalisiert (*Variable: Soziales Anerkennungsbedürfnis*). Die einzelnen Items der Skala stellen positive und negative Persönlichkeitseigenschaften dar, die sich ein Befragter zuschreiben oder

die er ablehnen kann. Die Zuschreibung positiver oder die Ablehnung negativer Eigenschaften werden jeweils mit dem Wert 1 und die Ablehnung positiver oder die Zuschreibung negativer Eigenschaften mit dem Wert 0 codiert. Da die Skala insgesamt 10 Items aufweist, entsprechen 10 erzielte Punkte einem maximalen Bedürfnis nach sozialer Anerkennung. Zudem wurde die von den Respondenten individuell wahrgenommene Sensitivität, die „Trait Desirability“, der einzelnen heiklen Fragen erfasst. Eine entsprechende Fragebatterie wurde in Anlehnung an einer erstmalig von Edwards (1953) vorgeschlagenen Operationalisierung konstruiert (vgl. auch Wyner 1980; Nicotera 1996; Huang et al. 1998). Die Respondenten wurden aufgefordert anzugeben, wie ihrer Meinung nach eine Person von der Gesellschaft bewertet wird, die das zuvor erfragte Verhalten aufweist. Die Antworten wurden anhand einer siebenstufigen Skala mit den Polen +3 (positiv bewertet) und -3 (negativ bewertet) abgegeben (*Variable: soziale Erwünschtheit des Verhaltens*). Zudem wird angenommen, dass mit zunehmender relativer Häufigkeit des Verhaltens im sozialen Umfeld eines Befragten der Sensitivitätsgrad einer heiklen Frage sinkt und somit mehr sozial unerwünschte Antworten zu erwarten sind (Johnson / van de Vijver 2002; Coutts / Jan 2008). Die Befragten wurden deshalb darum gebeten, auf einer vierstufigen Skala mit den Codierungen 1=„keiner“, 2=„manche“, 3=„viele“, 4=„alle“ einzuschätzen, wie viele Personen aus ihrem Bekanntenkreis das erfragte Verhalten aufweisen. Hierdurch können Rückschlüsse auf die Normen im sozialen Umfeld des Befragten gezogen werden (*Variable: wahrgenommene soziale Norm*). Um Antwortverzerrungen aufgrund fehlender Diskretion während der Antwortsituation zu kontrollieren, wurden die Befragten zudem aufgefordert, anhand einer fünfstufigen Skala mit den Polen 1=„überhaupt nicht vertrauen“ und 5=„voll vertrauen“ ihr Vertrauen in den Datenschutz bei dieser Befragung abzustufen (*Variable: Vertrauen in den Datenschutz*). Auch wurde nach der Anwesenheit dritter Personen (*Variable: Anwesenheit Dritter*) während des Ausfüllens des Fragebogens gefragt (vgl. Aquilino et al. 2000; Hartmann 1995; Reuband 1987, 1992). Die Antworten zu dieser Frage wurden mit 1=„Ja“ und 0=„Nein“ codiert.

Die Hypothesen 1-3 werden mit einem experimentellen Online-Survey überprüft. Diese Form der selbstadministrierten Erhebungsmethode gewinnt in jüngster Zeit zunehmend an Bedeutung (Couper / Coutts 2006). Die Rekrutierung von Umfrageteilnehmern erfolgte über das Online-Access-Panel *Sozioland* der Respondi AG.¹⁶ Auf die Möglichkeit, sich an dem

16 Fehlende Diskretion kann in Online-Surveys prinzipiell einen Anreiz für die Befragten darstellen, sozial unerwünschte Antworten zu vermeiden (Coutts / Jan 2008: 12): „Our inability to track response to the questionnaire means that participation in the study was nominally anonymous. However, although we do not know the identities of the survey participants, respondents may have had reason to believe that such knowledge was possible. Their identities are known to the panel organizers and they may have feared that their answers would be connected to those identities by ‘Socioland’.“ Um am Sozioland Panel teilzunehmen, müssen die potentiellen Befragten der Respondi AG sowohl Namen als auch Adressen mitteilen.

Access-Panel¹⁷ zu beteiligen (was mit einem Eintrag in eine Datenbank verbunden ist), werden potentielle Umfrageteilnehmer durch die Platzierung entsprechender Weblinks auf häufig besuchten Internetseiten aufmerksam gemacht und in das Access-Panel rekrutiert.¹⁸ Das zentrale Ziel der vorliegenden Untersuchung lautet kausale Effekte von experimentellen Stimuli auf eine abhängige Variable zu untersuchen. Die dabei erforderliche hohe interne Validität kann durch ein randomisiertes Experimentaldesign sichergestellt werden.¹⁹

Die Einladung zur Teilnahme an unserem experimentellen Survey erfolgte durch die Versendung von E-Mails an 10.000 zufällig ausgewählte Personen des Access-Panels.²⁰ Der Einladung folgten 1253 Teilnehmer, von denen N=1176 die Umfrage vollständig beendeten. Es wurden nur Personen im Alter ab 18 Jahren in die Analyse einbezogen. Ohne die 21 ausgeschlossenen Fälle von Personen unter 18 Jahren wurden schließlich Daten auf Grundlage der folgenden vier Versuchsgruppen analysiert: Gruppe 1: Direkte Formulierung (n=329, Kontrollgruppe) / Gruppe 2: Entschuldigende Formulierung (n=275) / Gruppe 3: Permissiver Kontext (n=285) / Gruppe 4: Restriktiver Kontext (n=266).

4 Empirische Ergebnisse

Einen Überblick über die Verteilungen der Antworten auf die sensitiven Fragen in den 4 Versuchsgruppen stellt Tabelle 1 dar. Um die Hypothesen 1-3 empirisch zu überprüfen wurden paarweise Pearson- χ^2 -Unabhängigkeitstests durchgeführt, wobei jeweils kein Unterschied in

-
- 17 Zu den Vorzügen von Access-Panels kann der relativ schnelle und ökonomische Aufbau eines Probandenpools gezählt werden, der sich zu einer Teilnahme an weiteren Befragungen bereit erklärt. Weitere Vorteile gegenüber traditionellen Umfragen sind u.a. die einfachere Ziehung von Zufallsstichproben aus dem Access-Panel Probandenpool sowie die Möglichkeit, potentielle Befragte über das Internet zu erreichen. Zu den bekannten Problemen des Access-Panel-Ansatzes lassen sich vor allem die beim Aufbau des Access-Panels auftretenden Selbstselektionsprozesse zählen (Couper / Couits 2006). Befragungen auf der Basis einer zufällig aus dem Access-Panel gezogenen Teilstichprobe sind insbesondere dann problematisch, wenn das primäre Untersuchungsziel die Schätzung von Prävalenzraten und die Verallgemeinerung auf eine größere Grundgesamtheit ist. So kann in den meisten Fällen die Stichprobenzusammensetzung auf der Grundlage eines Access-Panels nicht als „repräsentativ“ für die Bevölkerung Deutschlands betrachtet werden. Diese Problematik steht jedoch nicht im Vordergrund, wenn das zentrale Ziel einer Untersuchung lautet kausale Effekte von experimentellen Stimuli auf eine abhängige Variable zu untersuchen. In diesem Fall kann die erforderliche hohe interne Validität durch ein randomisiertes Experimentaldesign sichergestellt werden. Die implizite Annahme hierbei lautet, dass die Ausfallmechanismen bzw. Selbstselektionsprozesse in allen Experimentalgruppen identisch sind und sich nicht auf die interessierenden Unterschiede zwischen den Gruppen auswirken.
- 18 Access-Panels haben in der jüngsten Zeit zunehmend Verbreitung erfahren. Ein weiteres Beispiel für ein Access Panel ist die Dauerstichprobe (DSP) befragungsbereiter Haushalte auf Basis des Mikrozensus. Der vom statistischen Bundesamt organisierte und von den statistischen Landesämtern jährlich durchgeführte Mikrozensus befragt eine 1%-Zufallsstichprobe der Bevölkerung Deutschlands. Die DSP setzt sich aus Haushalten zusammen, die freiwillig ihre Teilnahmebereitschaft an Umfragen der amtlichen Statistik erklärt haben. Die Haushalte rekrutieren sich freiwillig aus dem jeweils jährlich ausscheidenden Rotationsviertel des Mikrozensus.
- 19 Das Prinzip der Randomisierung stellt sicher, dass es keine systematischen Unterschiede zwischen den Experimentalgruppen gibt und kontrolliert zudem neben bekannten Einflussfaktoren auch den Einfluss von unbekanntem Drittvariablen. Unterschiede in den abhängigen Variablen (hier: die selbstberichteten sozial erwünschten bzw. unerwünschten Verhaltensweisen) zwischen den Gruppen können somit kausal auf die Wirkung der experimentellen Stimuli (hier: Wording bzw. Fragekontext) zurückgeführt werden.
- 20 Den Angaben der Respondi AG zufolge besteht das Access-Panel aus ca. 50.000 Teilnehmern (Stand: 30.9.2008). Die Teilnehmer sind überwiegend weiblich (56.6%), jung (63.66% im Alter bis 29 Jahre) und gebildet (58.72% mit Abitur oder Fachabitur). Aus diesem Probandenpool werden über die Sozialand-Plattform regelmäßig Teilstichproben für Umfragen zu verschiedenen Themen (aus den Bereichen Kultur, Politik, Sport und Gesellschaft) gezogen.

den Prävalenzraten zwischen Experimental- und Kontrollgruppe als Nullhypothese angenommen wurde (Ergebnisse der Signifikanztests werden nicht berichtet). Es zeigen sich generell keine eindeutigen und konsistenten Unterschiede in den Prävalenzraten zwischen den Experimentalgruppen und der Kontrollgruppe in die theoretisch vorhergesagte Richtung.²¹

Tabelle 1: Verteilung sozial unerwünschter Antworten in Abhängigkeit von Versuchsgruppe und sensitivem Item

Sensitive Frage		Direkte Formulierung	Entschuldig. Formulierung	Permissiver Kontext	Restriktiver Kontext
Wahlbeteiligung	Nein	38 (12.97%)	21 (8.61%)	26 (10.66%)	15 (6.58%)
	Ja	255 (87.03%)	223 (91.39%)	218 (89.34%)	213 (93.42%)
Sexuelle Untreue	Nein	256 (78.05%)	228 (82.91%)	221 (77.82%)	222 (83.46%)
	Ja	72 (21.95%)	47 (17.09%)	63 (22.18%)	44 (16.54%)
Fahren unter Alkoholeinfluss	Nein	233 (71.04%)	192 (70.07%)	209 (73.33%)	187 (70.3%)
	Ja	95 (28.96%)	82 (29.93%)	76 (26.67%)	79 (29.7%)
Einnahme von Antidepressiva	Nein	304 (92.97%)	245 (89.42%)	263 (92.28%)	251 (94.36%)
	Ja	23 (7.03%)	29 (10.58%)	22 (7.72%)	15 (5.64%)

Abhängige Variablen sind jeweils die sensitiven Items in den Zeilen. Bedingte Häufigkeitsverteilungen sozial unerwünschter Antworten nach Experimentalgruppe und Item (mit spaltenbezogener Prozentuierung), sozial unerwünschtes Antwortverhalten ist fett markiert.

Die in den Hypothesen 1-3 postulierten Zusammenhänge lassen sich somit in den bivariaten Analysen empirisch nicht bestätigen. Um differenziertere Analysen durchzuführen und den Einfluss von Drittvariablen zu untersuchen, wurden in einem nächsten Schritt multiple logistische Regressionsmodelle geschätzt. Die abhängigen Variablen stellen jeweils dichotome Items mit den Ausprägungen „ja“ oder „nein“ dar. Hierbei wurden bei der Frage nach der Wahlbeteiligung nur Personen mit deutscher Staatsangehörigkeit berücksichtigt, die 18 Jahre oder älter waren. Da die letzte Bundestagswahl im Jahr 2005 stattfand, konnten Personen, die zum Zeitpunkt der Befragung (2008) unter 21 Jahre alt waren, an der Wahl nicht teilnehmen. Ausgeschlossen wurden infolge dessen 166 Fälle.²²

Die Zufallsvariable, anhand derer die Respondenten zu Kontroll- und Versuchsgruppen randomisiert wurden, stellt die Operationalisierung der Hypothesen 1-3 dar. Da diese Variable ein kategoriales Merkmal beschreibt, wurde sie zur Gewährleistung der Interpretierbarkeit

21 Zudem sind 10 von insgesamt 12 Tests nicht signifikant.

22 Aufgrund eines geringen Anteils an Item-Nonresponse sind leichte Variationen der Fallzahlen zwischen den Modellen zu beobachten: Sexuelle Untreue n=1127 / Fahren unter Alkoholeinfluss n=1134 / Einnahme von Antidepressiva n=1133. Es ist nicht davon auszugehen, dass sich diese wenigen Ausfälle systematisch zwischen den Experimentalgruppen unterscheiden und die interessierenden Differenzen zwischen den Gruppen systematisch verzerren.

jeweils für alle vier Ausprägungen in Dummy-Variablen recodiert. Damit entspricht jeweils eine Dummy-Variable einer Experimentalgruppe. Als Referenzkategorie wurde die Dummy-Variable der Kontrollgruppe „direkte Formulierung“ gewählt. Wording- und Kontexteffekte sind somit im Verhältnis zu dieser Referenzkategorie zu interpretieren. Die vier logistischen Regressionsmodelle wurden jeweils mit einem identischen Set von Kontrollvariablen geschätzt (vgl. Beschreibung in Abschnitt 3). Zudem wurden weitere Kontrollvariablen in die Regressionsmodelle aufgenommen. Es handelt sich hierbei um die soziodemographischen Variablen *Geschlecht*, *Fester Partner*, und *Alter* sowie die Variable *generalisiertes Vertrauen* der Befragten. So konnte in einer empirischen Studie zu Fremdenfeindlichkeit und Antisemitismus (vgl. Krumpal 2010) ein klarer Zusammenhang zwischen generalisiertem Vertrauen und sozial unerwünschten Antworten gezeigt werden.

Die McFadden-Pseudo- R^2 -Werte in Tabelle 2 deuten jeweils auf einen akzeptablen bis guten Gesamt-Fit der logistischen Regressionsmodelle hin.²³ Zudem deuten in allen vier Modellen die ermittelten LR- χ^2 -Werte auf einen signifikanten Modell-Fit hin ($p < 0.01$).

Laut den Hypothesen 1-3 werden folgende Odds-Ratios erwartet: Da die Beteiligung an der letzten Bundestagswahl als sozial erwünscht betrachtet wird, ist gemäß Hypothese 1 für eine Formulierung, die die Nicht-Beteiligung an der letzten Wahl entschuldigt, ein Odds-Ratio < 1 (Kodierung der abhängigen Variable: „ja“=1 / „nein“=0) zu erwarten. Sexuelle Untreue in der Partnerschaft, das Steuern eines Kraftfahrzeugs unter starkem Alkoholeinfluss und die Einnahme von Antidepressiva werden dagegen als sozial unerwünschte Verhaltensweisen betrachtet. Für entschuldigende Formulierungen sind folglich Odds-Ratios > 1 (Kodierung der abhängigen Variable: „ja“=1 / „nein“=0) zu erwarten. Die gleichen Erwartungen gelten für die Überprüfung von Hypothese 2. Ein permissiver Fragekontext verringert im Vergleich zu einer direkten Frageformulierung die kognitive Dissonanz, die durch eine sensitive Frage induziert wird. Deshalb sind gemäß den Annahmen über die Sensitivitätsrichtung der einzelnen Fragen für die Frage nach der Wahlbeteiligung ein Odds-Ratio < 1 und für alle anderen Fragen Odds-Ratios > 1 zu erwarten. Da schließlich laut Hypothese 3 ein restriktiver Fragekontext die bestehende kognitive Dissonanz im Vergleich zu einer direkten Frageformulierung verstärkt, sind für die Frage nach der Wahlbeteiligung ein Odds-Ratio > 1 und für die anderen sensitiven Fragen Odds-Ratios < 1 zu erwarten.

23 Angesichts der Tatsache, dass die Berechnung vieler Pseudo- R^2 -Anpassungsmaße generell in niedrigeren Werten resultiert, als dies für die R^2 -Werte linearer Regressionsmodelle gilt (Hosmer / Lemeshow 2000), sind die berechneten Modellanpassungen als akzeptabel zu betrachten.

Tabelle 2: Wording- und Kontexteffekte bei sensitiven Fragen – Logistische Regressionen, Odds Ratios

	Wahl- beteiligung	Sexuelle Untreue	Fahren unter Alkohol- einfluss	Einnahme von Anti- depressiva
Entschuldig. Formulierung (Ref.: Direkte Formulierung)	1.63 (1.52)	0.79 (-1.04)	1.12 (0.52)	1.72 ⁺ (1.76)
Permissiver Kontext (Ref.: Direkte Formulierung)	1.36 (1.0)	0.91 (-0.42)	0.95 (-0.23)	1.19 (0.53)
Restriktiver Kontext (Ref.: Direkte Formulierung)	1.85 ⁺ (1.74)	0.66 ⁺ (-1.75)	1.33 (1.28)	0.85 (-0.45)
Soziales Anerkennungs- Bedürfnis	1.09 (1.61)	0.9** (-2.77)	0.91* (-2.53)	0.93 (-1.29)
Soziale Erwünschtheit des Verhaltens	1.09 (0.82)	1.13 ⁺ (1.66)	1.23** (3.03)	1.02 (0.19)
Wahrgenommene Soziale Norm	5.67** (8.35)	3.57** (7.28)	4.69** (10.08)	3.53** (5.94)
Vertrauen in den Datenschutz	1.12 (0.79)	0.85 ⁺ (-1.67)	0.98 (-0.2)	0.8 ⁺ (-1.71)
Anwesenheit Dritter (Dritte anw=1; Dritte n. anw.=0)	0.4 (-1.63)	0.28 ⁺ (-1.86)	1.28 (0.55)	0.36 (-0.99)
Generalisiertes Vertrauen (Vertrauen=1; kein Vertrauen=0)	1.15 (0.55)	1.4* (1.99)	0.86 (-0.92)	0.62 ⁺ (-1.95)
Geschlecht (Mann=1; Frau=0)	1.39 (1.29)	1.0 (-0.03)	3.43** (7.55)	0.69 (-1.47)
Fester Partner (Partner=1; kein Partner=0)	2.15** (3.13)	1.9** (3.35)	0.81 (-1.2)	0.96 (-0.18)
Alter	1.01 (0.82)	1.03** (4.39)	1.05** (7.19)	1.03** (3.72)
<i>N</i>	989	1127	1134	1133
<i>LR</i> – χ^2	122.6**	155.07**	348.51**	75.53**
<i>McFadden</i> – <i>R</i> ²	0.1946	0.1397	0.2558	0.1211

Abhängige Variablen sind jeweils die dichotomen sensitiven Items (Antwortmöglichkeiten: 1= „ja“ / 0= „nein“) in den Spalten. Sozial unerwünschte Antworten: „nein“ für Wahlbeteiligung / „ja“ für sexuelle

Untreue, Fahren unter Alkoholeinfluss und Einnahme von Antidepressiva / z-Statistiken in Klammern unterhalb der Odds Ratios. Empirisches Signifikanzniveau (p-Werte): ⁺ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Die Ergebnisse in Tabelle 2 zeigen insgesamt keine eindeutigen und konsistenten Effekte entschuldigender Frageformulierungen (bzw. eines permissiven oder restriktiven Fragekontextes) auf die Chancen, ein sozial unerwünschtes bzw. privates Verhalten zu berichten. Die Ergebnisse der bivariaten Analysen können somit im Rahmen einer multiplen logistischen Regressionsanalyse reproduziert werden.

Für die Kontrollvariablen *soziales Anerkennungsbedürfnis*, *soziale Erwünschtheit des Verhaltens*, und *wahrgenommene soziale Norm* zeigen alle Odds-Ratios ausnahmslos über alle sensitiven Fragen hinweg in die theoretisch erwartete Richtung und sind in der Mehrzahl der Fälle auf den konventionellen Niveaus signifikant (8 von 12 Tests): So steigen, wie erwartet, bei zunehmendem Bedürfnis nach sozialer Anerkennung die Chancen, eine sozial erwünschte Antwort „ja“ zu geben (Wahlbeteiligung, $OR=1.09$, nicht signifikant), bzw. sinken die Chancen, eine sozial unerwünschte Antwort „ja“ zu geben (sexuelle Untreue, $OR=0.90$, $p < 0.01$ / Fahren unter Alkoholeinfluss, $OR=0.91$, $p < 0.05$ / Einnahme von Antidepressiva / $OR=0.93$, nicht signifikant). Erwartungsgemäß steigen zudem mit zunehmender sozialer Erwünschtheit des erfragten Verhaltens die Chancen, das Verhalten zu berichten und die entsprechende Frage mit „ja“ zu beantworten (Wahlbeteiligung, $OR=1.09$, nicht signifikant / sexuelle Untreue, $OR=1.13$, $p < 0.1$ / Fahren unter Alkoholeinfluss, $OR=1.23$, $p < 0.01$ / Einnahme von Antidepressiva / $OR=1.02$, nicht signifikant). Die Variable *wahrgenommene soziale Norm* zeigt mit Abstand den stärksten und konsistentesten Effekt auf die Bereitschaft des Befragten, potentiell peinliches bzw. sozial unerwünschtes Verhalten zu berichten: Mit zunehmendem wahrgenommenen Anteil des erfragten Verhaltens im sozialen Umfeld des Befragten steigen die Chancen, das entsprechende Verhalten zu berichten, stark und auch signifikant an (Wahlbeteiligung, $OR=5.67$, $p < 0.01$ / sexuelle Untreue, $OR=3.57$, $p < 0.01$ / Fahren unter Alkoholeinfluss, $OR=4.69$, $p < 0.01$ / Einnahme von Antidepressiva / $OR=3.53$, $p < 0.01$).

Die Kontrollvariablen *Anwesenheit Dritter*, *Vertrauen in den Datenschutz*, und *generalisiertes Vertrauen* zeigen dagegen weniger konsistente Effekte. So zeigt die Anwesenheit dritter Personen während der Beantwortung des Online-Fragebogens in drei von vier Fällen einen Effekt in die erwartete Richtung. Demnach werden in Anwesenheit dritter Personen bei den Fragen zu Wahlbeteiligung, sexueller Untreue, und Einnahme von Antidepressiva mehr sozial erwünschte Antworten gegeben. Hierbei wirkt sich lediglich bei der Variable *sexuelle Untreue* die Anwesenheit Dritter während der Datenerhebung signifikant auf das Ausmaß sozial erwünschter Antworten aus ($OR=0.28$ / $p < 0.1$). Dagegen zeigen die Odds-Ratios der Variable *Vertrauen in den Datenschutz* konsistent in die nicht erwartete Richtung (in zwei Fällen signifikant): So sinken bei den Fragen nach sexueller Untreue in der Partnerschaft und der Einnahme von Antidepressiva mit zunehmendem Vertrauen in den Datenschutz signifikant die Chancen, eine sozial unerwünschte „ja“-Antwort zu geben (sexuelle Untreue: $OR=0.85$ /

$p < 0.1$ / Einnahme von Antidepressiva $OR = 0.8$ / $p < 0.1$).²⁴ Schließlich zeigen sich auch beim *generalisierten Vertrauen* und den soziodemographischen Kontrollvariablen keine konsistenten und klar interpretierbaren Zusammenhänge mit dem sozial erwünschten Antwortverhalten der Befragten.

Im theoretischen Teil der Arbeit wurde als notwendige Bedingung für das Entstehen kognitiver Dissonanz angenommen, dass vom Respondenten eine soziale Norm geteilt wird, die dem erfragten Verhalten widerspricht. Zudem wurde angenommen, dass sich Unterschiede im Antwortverhalten durch die Addition *konsonanter* (entschuldigende Formulierung, permissiver Kontext) bzw. *dissonanter* Kognitionen (restriktiver Kontext) erklären lassen. In diesem Zusammenhang ist ein modifizierter Wirkungsmechanismus denkbar, wonach die experimentellen Stimuli nicht direkt, durch die Addition *konsonanter* bzw. *dissonanter* Kognitionen, sondern vielmehr indirekt, über die wahrgenommene Normgültigkeit bzw. die sozialen Erwünschtheitswahrnehmungen vermittelt, auf das Antwortverhalten wirken. Im Falle der wahrgenommenen Normgültigkeit wird angenommen, dass bei einer entschuldigenden Frageformulierung bzw. einem permissivem Kontext die Wahrscheinlichkeit für eine positive Normwahrnehmung hinsichtlich des erfragten Verhaltens steigt und bei einem restriktiven Kontext dagegen sinkt. Im Falle der sozialen Erwünschtheitswahrnehmungen wird angenommen, dass die experimentellen Variationen in den Formulierungen und im Kontext der sensitiven Fragen die soziale Unerwünschtheit des Verhaltens verringern (entschuldigende Formulierung, permissiver Kontext), bzw. erhöhen (restriktiver Kontext). Diese modifizierten Annahmen wurden anhand von einfachen OLS-Regressionen bzw. Heterogeneous-Choice-Modellen (Keele / Park 2006; Williams 2007) empirisch überprüft.

-
- 24 In diesem Zusammenhang wäre ein Interaktionseffekt zwischen den experimentellen Stimuli und der Vertrauensvariable denkbar. So könnte man annehmen, dass Effekte der Formulierung und des Kontextes auf das Antwortverhalten vor allem bei Befragten aktiviert werden, die kein bzw. nur ein geringes Vertrauen in den Datenschutz haben. Um diese Annahme zu prüfen, wurden die Befragten in drei Untergruppen aufgeteilt: Gruppe 1: kein bzw. wenig Vertrauen in den Datenschutz, Gruppe 2: bezüglich des Vertrauens in den Datenschutz neutrale Respondenten, Gruppe 3: hohes bzw. volles Vertrauen in den Datenschutz. Für die drei Gruppen wurden anschließend bivariate logistische Regressionen berechnet (unabhängige Variable=Experimentalbedingung; abhängige Variable=die dichotomen sensitiven Items analog zur Tabelle 2). Demnach zeigt sich kein klarer Interaktionseffekt zwischen der Formulierung bzw. des Kontextes und der Vertrauensvariable auf das Antwortverhalten. Nur im Zusammenhang mit Fahren unter Alkoholeinfluss erhöht der permissive Kontext signifikant die Wahrscheinlichkeit von sozial unerwünschten Antworten bei Befragten, die kein bzw. wenig Vertrauen in den Datenschutz aufweisen ($OR = 3.56$ / $p < 0.01$). Diese zusätzlichen Analysen können bei Bedarf von den Autoren angefordert werden.

Tabelle 3: Wording und Kontexteffekte auf die Wahrnehmung der Verhaltensnorm: Heterogeneous Choice – Modelle, Odds Ratios

Wahrgenommene soziale Norm hinsichtlich...	Wahlbeteiligung	sexueller Untreue	des Fahrens unter Alkoholeinfluss	der Einnahme von Antidepressiva
Entschuldig. Formulierung (Ref.: Direkte Formulierung)	1.01 (0.08)	0.9 (-0.62)	0.94 (-0.4)	1.05 (0.3)
Permissiver Kontext (Ref.: Direkte Formulierung)	1.02 (0.12)	1.33 (1.63)	0.82 (-1.24)	0.98 (-0.1)
Restriktiver Kontext (Ref.: Direkte Formulierung)	1.36 (1.64)	1.28 (1.39)	0.78 (-1.5)	0.87 (-0.86)
<i>N</i>	1005	1145	1151	1149
<i>LR-χ^2</i>	3.47	6.76 ⁺	2.93	1.33
<i>McFadden – R²</i>	0.002	0.004	0.001	0.001

Die abhängige Variable „wahrgenommene soziale Norm“ besteht aus einer vierstufigen Skala mit den Antwortmöglichkeiten 1= „keiner“ 2= „manche“ 3= „viele“ 4= „alle“ / z-Statistiken in Klammern unterhalb der Odds Ratios. Empirisches Signifikanzniveau (p-Werte): ⁺ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$.

Tabelle 4: Wording- und Kontexteffekte auf die soziale Erwünschtheit des Verhaltens: Bivariate OLS-Regressionen

Soziale Erwünschtheit...	der Wahl- beteiligung	sexueller Untreue	Fahrens unter Alkohol- einfluss	der Einnahme von Anti- depressiva
Entschuldig. Formulierung (Ref.: Direkte Formulierung)	-0.17 ⁺ (-1.70)	0.09 (0.98)	0.17 ⁺ (1.77)	0.02 (0.26)
Permissiver Kontext (Ref.: Direkte Formulierung)	-0.2 ⁺ (-1.94)	0.08 (0.95)	0.17 ⁺ (1.73)	0.1 (1.24)
Restriktiver Kontext (Ref.: Direkte Formulierung)	-0.14 (-1.38)	0.16 ⁺ (1.83)	0.13 (1.33)	0.12 (1.49)
Konstante	1.42** (20.73)	-1.71** (-28.63)	-1.71** (-26.29)	-0.58** (-10.43)
<i>N</i>	1004	1152	1153	1153
<i>F</i> -Statistik	1.57	1.12	1.43	1.03
<i>R</i> ²	0.05	0.03	0.004	0.003

Abhängige Variable ist jeweils die „soziale Erwünschtheit des Verhaltens“ (Siebenstufige Antwortskala mit den Polen -3 „negativ bewertet“ und +3 „positiv bewertet“) / t-Statistiken in Klammern unterhalb der Regressionskoeffizienten. Empirisches Signifikanzniveau (p-Werte): ⁺p<0.1, *p<0.05, **p<0.01.

Wie aus den Tabellen 3 und 4 ersichtlich wird, haben die experimentellen Variationen der Frageformulierung bzw. des Fragekontextes keinen signifikanten Einfluss auf die Variable wahrgenommene soziale Norm bzw. zeigen nur vereinzelt marginal signifikante Effekte auf die Variable soziale Erwünschtheit des Verhaltens. Die sozialen Norm- und Erwünschtheitswahrnehmungen der Befragten sind demnach recht robust gegenüber Manipulationen der Frageformulierung bzw. des Fragekontextes und beeinflussen das Antwortverhalten unabhängig von diesen.

5 Zusammenfassung und Diskussion

Als Fazit zeigt sich, dass die empirische Evidenz zu Wording- und Kontexteffekten bei sensitiven Fragen weiterhin als uneinheitlich zu betrachten ist. Es zeigen sich generell keine nennenswerten Auswirkungen von *entschuldigenden Frageformulierungen* auf die Bereitschaft der Befragten, mehr sozial unerwünschtes Verhalten zu berichten. Die Ergebnisse unseres experimentellen Surveys stehen in Übereinstimmung mit den Ergebnissen von Presser (1990) und Catania et al. (1996). Auch hier finden sich keine konsistenten Kontext- bzw. Wording-Effekte auf Antworten zu Fragen nach sozial unerwünschtem Verhalten. Ebenso deuten auch die Befunde von Holtgraves et al. (1997) auf lediglich schwache und wenig konsistente Effekte hin. Mit Hinblick auf sozial erwünschtes Verhalten widersprechen unsere Ergebnisse Belli et al. (2006). In dieser Studie konnten signifikante Effekte einer entschuldigenden Frageformulierung auf die Angaben zur Wahlbeteiligung gezeigt werden. Somit sind die praktischen Empfehlungen in Lehrbüchern der Survey Methodology (Sudman / Bradburn 1982; Fowler 1995; Groves et al. 2004), sensitive Fragen so zu formulieren, dass das erfragte Verhalten entschuldigt wird, mit Skepsis zu betrachten.

Dagegen wirkt sich bei allen sensitiven Fragen die wahrgenommene relative Häufigkeit des erfragten Verhaltens im sozialen Umfeld der Respondenten signifikant auf das Antwortverhalten aus (*wahrgenommene soziale Norm*). Die Wirksamkeit einer solchen Norm, die in der vorliegenden Arbeit im Sinne einer empirischen Verhaltensregelmäßigkeit im sozialen Umfeld des Befragten interpretiert wird (vgl. hierzu auch die Deutung des Normbegriffs als „behavioral regularities“ bei Opp 2001 a), ist konsistent mit den Annahmen des „Hypocrisy“-Paradigmas: Demnach würde sich mit zunehmend höherer Einschätzung der relativen Verhaltenshäufigkeit im sozialen Umfeld die Stärke der induzierten kognitiven Dissonanz bei den Befragten, bei denen das sozial unerwünschte Verhalten tatsächlich vorliegt, verringern. Entsprechend liegt eine theoretische Fundierung der empirisch beobachteten Effekte vor, wonach mit zunehmendem Verhaltensanteil im sozialen Umfeld die Chancen stark ansteigen, sexuelle Untreue, das Fahren eines Kraftfahrzeugs unter Alkoholeinfluss, und auch die Einnahme von Antidepressiva zu berichten.²⁵ Gemäß dem „Hypocrisy“-Paradigma ist zudem kognitive Dissonanz zu erwarten, wenn ein Respondent die letzte Bundestagswahl *nicht* besucht hat, aber in seinem sozialen Umfeld eine hohe Wahlbeteiligung wahrnimmt. Das vorliegende Ergebnis, mit zunehmender Wahlbeteiligung im sozialen Umfeld die Frage nach der Wahlbeteiligung eher zu bejahen, kann daher als Versuch des Befragten interpretiert werden, kognitive Dissonanz zu verringern.²⁶

Zukünftige Forschung könnte die Auswirkungen von Wording- und Kontexteffekten auf sozial erwünschtes Antwortverhalten in unterschiedlichen Befragungsmodi vergleichend untersuchen: In der vorliegenden Arbeit wurde angenommen, dass Befragte in Online Surveys unter bestimmten Bedingungen dazu neigen können, sozial unerwünschte Antworten zu vermeiden. Dies kann beispielsweise dann der Fall sein, wenn die Respondenten annehmen, dass ihre Antworten im Prinzip mit ihren Identitäten in Verbindung gebracht werden können (Coutts / Jan 2008: 12). So liefert die neuere Validierungsstudie von Kreuter et al. (2008) klare Belege für sozial erwünschtes Antwortverhalten in Online Surveys. Auch in unserer Studie zeigen sich eindeutige Zusammenhänge zwischen der Neigung zu sozial erwünschten Antworten einerseits, und der angenommenen Normgültigkeit, dem sozialen Anerkennungsbedürfnis bzw. den Erwünschtheitswahrnehmungen der Befragten andererseits. Allerdings könnte man argumentieren, dass im Vergleich zu einem Verfahren mit einem Interviewer, die Aktivierung von Effekten sozialer Erwünschtheit in selbstadministrierten Befragungsmodi vergleichsweise geringer sei (Tourangeau / Yan 2007; Kreuter et al. 2008). Ein solcher Modus-effekt wäre eine mögliche Erklärung der nur schwachen und uneinheitlichen Wording- und Kontexteffekte auf das Antwortverhalten der Befragten. Als fruchtbare Anknüpfung an unseren Beitrag sollten deshalb in zukünftigen Studien erweiterte Forschungsdesigns entwickelt

- 25 Die theoretische Rekonstruktion der Wirksamkeit der wahrgenommenen relativen Verhaltenshäufigkeit bzw. des wahrgenommenen Verhaltensanteils im Rahmen des „Hypocrisy“-Paradigmas setzt voraus, dass (1) dieser Anteil als soziale Norm (im Sinne von empirischen Regelmäßigkeiten und beobachteten Verhaltensstendenzen im sozialen Umfeld des Befragten) gedeutet werden kann, und (2) diese Norm vom Befragten internalisiert wurde (in Form von Kognitionen was in der eigenen Lebenswelt als „normal“ wahrgenommen wird).
- 26 Allerdings ist eine weitere Interpretationsmöglichkeit nicht zu übersehen: Die relative Häufigkeitseinschätzung im Sinne eines wahrgenommenen Anteils stellt einen Indikator für die Norm dar, die ein Befragter in seinem sozialen Umfeld wahrnimmt. Zwischen sozialen Normen und tatsächlichen Verhalten liegt ein positiver Zusammenhang vor (vgl. Rauhut / Krumpal 2008). Deshalb kann angenommen werden, dass mit zunehmend höherer Einschätzung der relativen Verhaltenshäufigkeit im sozialen Umfeld die Wahrscheinlichkeit zunimmt, dass das Verhalten bei einem Befragten tatsächlich vorliegt. Auf Aggregatdatenebene lassen sich demnach mit zunehmender Stärke der sozialen Norm auch höhere Anteile des tatsächlichen Verhaltens erwarten. Dies könnte die recht starken Effekte der Variable wahrgenommene soziale Norm bei allen vier sensitiven Fragen teilweise erklären. Empirisch lässt sich die relative Stärke dieses Effektes allerdings nur durch eine Studie mit externen Validierungsdaten, bei denen das tatsächliche Verhalten bekannt ist, genauer einschätzen.

werden, die sowohl das Wording (bzw. den Kontext) als auch den Befragungsmodus (selbst-administriert versus interviewer-administriert) experimentell variieren, um einen denkbaren Interaktionseffekt weiterführend zu untersuchen. Speziell könnte die Hypothese untersucht werden, wonach die postulierten Wording- und Kontexteffekte nur in Anwesenheit eines Interviewers bzw. bei geringerer Anonymität der Interviewbedingungen auftreten. In einer möglichen Folgestudie wäre es schließlich auch sehr wünschenswert, die Antworten mit externen Daten zu validieren.

Literatur

- Aquilino, W. S. / D.L. Wright / A.J. Supple (2000): Response Effects due to Bystander Presence in CASI and Paper-and-Pencil Surveys of Drug Use and Alcohol Use, in: *Substance Use & Misuse* 35, S. 845-867.
- Aronson, E. (1999): Dissonance Hypocrisy and the Self-Concept, in: E. Harmon-Jones / J. Mills (Hrsg.): *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology*. Washington D.C., S. 103-106.
- Aronson, E. / C.B. Fried / J. Stone (1991): Overcoming Denial and increasing the Intention to use Condoms through the Induction of Hypocrisy, in: *American Journal of Public Health* 81, S. 1636-1638.
- Barton, A.H. (1958): Asking the Embarrassing Question, in: *Public Opinion Quarterly* 22, S. 67-68.
- Becker, R. (2006): Selective Response to Questions on Delinquency, in: *Quality & Quantity* 40, S. 483-498.
- Becker, R. / R. Günther (2004): Selektives Antwortverhalten bei Fragen zum delinquenten Handeln – Eine empirische Studie über die Wirksamkeit der „sealed envelope technique“ bei selbstberichteter Delinquenz mit Daten des ALLBUS 2000, in: *ZUMA-Nachrichten* 54, S. 39-59.
- Belli, R.F. / S.E. Moore / J. VanHoewyk (2006): An experimental Comparison of Question Forms used to reduce Vote Overreporting, in: *Electoral Studies* 25, S. 751-759.
- Belli, R.F. / M.W. Traugott / M.N. Beckmann (2001): What leads to Voting Overreports? Contrasts of Overreporters to Validated Voters and Admitted Nonvoters in the American National Election Studies, in: *Journal of Official Statistics* 17, S. 479-498.
- Beyer, H. / I. Krumpal (2010): „Aber es gibt keine Antisemiten mehr“: Eine experimentelle Studie zur Kommunikationslatenz antisemitischer Einstellungen, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 62, S. 681-705.
- Blair, E. / S. Sudman / N.M. Bradburn / C. Stocking (1977): How to ask Questions about Drinking and Sex: Response Effects in Measuring Consumer Behavior, in: *Journal of Marketing Research* 14, S. 316-321.
- Cantril, H. / S.S. Wilks (1940): Problems and Techniques – Experiments in the Wording of Questions, in: *Public Opinion Quarterly* 4, S. 330-338.
- Catania, J.A. / D. Binson / J. Canchola / L.M. Pollack / W. Hauck (1996): Effects of Interviewer Gender, Interviewer Choice, and Item Wording on Responses to Questions concerning Sexual Behavior, in: *Public Opinion Quarterly* 60, S. 345-375.
- Couper, M.P. / E. Coultts (2006): Probleme und Chancen verschiedener Arten von Online-Erhebungen, in: A. Diekmann (Hrsg.), *Methoden der Sozialforschung. Sonderheft 44 der Kölner Zeitschrift für Soziologie und Sozialpsychologie*, S. 217-243.
- Coultts, E. / B. Jann (2008): Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT), ETH Zürich, Professuren für Soziologie, Arbeitspapier.
- Coultts, E. / B. Jann (2011): Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT), in: *Sociological Methods and Research* 40 / 1, S. 169-193.

- Coutts, E. / B. Jann / I. Krumpal / A.-F. Näher (2011): Plagiarism in Student Papers: Prevalence Estimates Using Special Techniques for Sensitive Questions, in: *Journal of Economics and Statistics* 231 / 5+6, S. 749-760.
- Crowne, D.P. / D. Marlowe (1960): A New Scale of Social Desirability Independent of Psychopathology, in: *Journal of Consulting Psychology* 24, S. 349-354.
- Diekmann, A. (2003): *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen*, 10. Aufl., Reinbek.
- Edwards, A.L. (1953): The Relationship between the judged Desirability of a Trait and the Probability that the Trait will be endorsed, in: *Journal of Applied Psychology* 2, S. 90-93.
- Esser, H. (1986): Können Befragte lügen? Zum Konzept des „wahren Wertes“ im Rahmen der handlungstheoretischen Erklärung von Situationseinflüssen bei der Befragung, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 38, S. 314-336.
- Festinger, L. (1957): *A Theory of Cognitive Dissonance*, Stanford.
- Fointiat, V. (2008): Being together in a Situation of induced Hypocrisy, in: *Current Research in Social Psychology* 13, S. 145-153.
- Fowler, F.J. Jr. (1995): *Improving Survey Questions. Design and Evaluation*, London.
- Frey, D. / A. Gaska (1993): Die Theorie der kognitiven Dissonanz, S. 274-324 in: Frey, D. / M. Irlé (Hrsg.): *Theorien der Sozialpsychologie. Band 1: Kognitive Theorien*, 2. Aufl., Bern.
- Fu, H. / J.E. Darroch / S.K. Henshaw / E. Kolb (1995): Measuring the Extent of Abortion Underreporting in the 1995 National Survey of Family Growth, in: *Family Planning Perspectives* 30, S. 128-133 & 138.
- Greifeneder, R. / H. Bless (2010): The fate of activated information in impression formation: Fluency of concept activation moderates the emergence of assimilation versus contrast, in: *British Journal of Social Psychology* 49, S. 405-414.
- Groves, R.M. / F.J. Fowler / M.P. Couper / J.M. Lepkowski / E. Singer / R. Tourangeau (2004): *Survey Methodology*, Hoboken.
- Hartmann, P. (1995): Response Behavior in Interview Settings of limited Privacy, in: *International Journal of Public Opinion Research* 7, S. 383-390.
- Holtgraves, T. (2004): Social Desirability and Self-Reports: Testing Models of Socially Desirable Responding, in: *Personality and Social Psychology Bulletin* 30, S. 161-172.
- Holtgraves, T. / J. Eck / B. Lasky (1997): Face Management, Question Wording and Social Desirability, in: *Journal of Applied Social Psychology* 27, S. 1650-1671.
- Hosmer, D.W. / S. Lemeshow (2000): *Applied Logistic Regression*, New York.
- Huang, C. / H. Liao / S.-H. Chang (1998): Social Desirability and the Clinical Self-Report Inventory: Methodological Reconsideration, in: *Journal of Clinical Psychology* 54, S. 517-528.
- Jann, B. / J. Jerke / I. Krumpal (2012): Asking Sensitive Questions Using the Crosswise Model: An Experimental Survey Measuring Plagiarism, in: *Public Opinion Quarterly* 76, S. 32-49.
- Johnson, T.J. / F.J. van de Vijver (2002): Social Desirability in Cross-Cultural Research, in: J. Harness / F.J. van de Vijver / P. Mohler (Hrsg.), *Cross-Cultural Survey Methods*, New York, S. 193-202.
- Keele, L. / D.K. Park (2006): Difficult choices: an evaluation of heterogeneous choice models, Annual Meeting of the American Political Science Association, Philadelphia 2006.
- Kreuter, F. / S. Presser / R. Tourangeau (2008): Social Desirability Bias in CATI, IVR, and Web Surveys: The Effects of Mode and Question Sensitivity, in: *Public Opinion Quarterly* 72 / 5, S. 847-865.
- Krumpal, I. (2010): *Sensitive Questions and Measurement Error: Using the Randomized Response Technique to Reduce Social Desirability Bias in CATI Surveys*, Universität Leipzig, Fakultät für Sozialwissenschaften und Philosophie, Dissertation.
- Krumpal, I. (2011): Determinants of Social Desirability Bias in Sensitive Surveys: A Literature Review, in: *Quality & Quantity* (Online First), doi: 10.1007/s11135-011-9640-9.

- Leak, G.K. / S. Fish (1989): Religious Orientation, Impression Management, and Self-Deception: Toward a Clarification of the Link between Religiosity and Social Desirability, in: *Journal of the Scientific Study of Religion* 28 / 3, S. 355-359.
- Nederhof, A.J. (1985): Methods of coping with social desirability bias: a review, in: *European Journal of Social Psychology* 15, S. 263-280.
- Nicotera, A.M. (1996): An Assessment of the Argumentativeness Scale for Social Desirability Bias, in: *Communication Reports* 9, S. 23-36.
- Opp, K.-D. (2001 a): Norms, in: P.B. Baltes / N.J. Smelser (Hrsg.), *International Encyclopedia of Social and Behavioral Sciences*, Amsterdam, S. 10714-10720.
- Opp, K.-D. (2001 b): Social Networks and the Emergence of Protest Norms, in: M. Hechter / K.-D. Opp (Hrsg.), *Social Norms*, New York, S. 234-273.
- Paulhus, D.L. (1984): Two-Component Models of Socially Desirable Responding, in: *Journal of Personality and Social Psychology* 46, S. 598-609.
- Paulhus, D.L. (2002): Socially 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*, Mahwah, S. 49-69.
- Presser, S. (1990): Can Changes in Context reduce Vote Overreporting in Surveys?, in: *Public Opinion Quarterly* 54, S. 586-593.
- Randall, D.M. / M.F. Fernandes (1991): The Social Desirability Response Bias in Ethics Research, in: *Journal of Business Ethics* 10, S. 805-817.
- Rauhut, H. / I. Krumpal (2008): Die Durchsetzung sozialer Normen in Low-Cost- und High-Cost-Situationen, in: *Zeitschrift für Soziologie* 37, S. 380-402.
- Reuband, K.H. (1987): Unerwünschte Dritte beim Interview: Erscheinungsformen und Folgen, in: *Zeitschrift für Soziologie* 16, S. 303-308.
- Reuband, K.H. (1992): On 3rd Persons in the Interview Situation and their Impact on Responses, in: *International Journal of Public Opinion Research* 4, S. 269-274.
- Schnell, R. / P.B. Hill / E. Esser (1999): *Methoden der empirischen Sozialforschung*, 6. Aufl., München – Wien.
- Schwarz, N. / H. Bless (1992): Scandals and the Public Trust in Politicians: Assimilation and Contrast Effects, in: *Personality and Social Psychology Bulletin* 18, S. 574-579.
- Schwarz, N. / H. Bless (2007): Mental Construal Processes: The Inclusion / Exclusion Model, in: D.A. Stapel / J. Suls (Hrsg.), *Assimilation and Contrast in Social Psychology*, Philadelphia, S. 119-141.
- Skarbak-Kozietulska, A. / P. Preisendörfer / F. Wolter (2012): Leugnen oder gestehen? Bestimmungsfaktoren wahrer Antworten in Befragungen, in: *Zeitschrift für Soziologie* 41, S. 5-23.
- Stocké, V. (2004): Entstehungsbedingungen von Antwortverzerrungen durch soziale Erwünschtheit. Ein Vergleich der Prognosen der Rational-Choice Theorie und des Modells der Frame-Selektion, in: *Zeitschrift für Soziologie* 33, S. 303-320.
- Stocké, V. (2007): Deutsche Kurzsкала des Bedürfnisses nach sozialer Anerkennung, in: A. Glöckner-Rist (Hrsg.), *ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente*, ZIS-Version 11.0, Bonn: GESIS.
- Stocké, V. / C. Hunkler (2007): Measures of Desirability Beliefs and their Validity as Indicators for Socially Desirable Responding, in: *Field Methods* 19, S. 313-336.
- Stone, J. / E. Aronson / A.L. Crain / M.P. Winslow / C.B. Fried (1994): Inducing Hypocrisy as a Means of Encouraging young Adults to use Condoms, in: *Personality and Social Psychology Bulletin* 20, S. 116-128.
- Stone, J. / A.W. Wiegand / J. Cooper / E. Aronson (1997): When Exemplification fails: Hypocrisy and the Motive for Self-Integrity, in: *Journal of Personality and Social Psychology* 72, S. 54-65.

- Strack, F. / N. Schwarz / M. Wänke (1991): Semantic and Pragmatic Aspects of Context Effects in Social and Psychological Research, in: *Social Cognition* 9, S. 111-125.
- Sudman, S. / N.M. Bradburn (1982): *Asking Questions: A Practical Guide to Questionnaire Design*, San Francisco.
- Sudman, S. / N.M. Bradburn / N. Schwarz (1996): *Thinking about Answers. The Application of Cognitive Processes to Survey Methodology*, San Francisco.
- Tourangeau, R. / K. Rasinski / N.M. Bradburn / R. D'Andrade (1989): Carry-over Effects in Attitude Surveys, in: *Public Opinion Quarterly* 53, S. 495-524.
- Tourangeau, R. / L.J. Rips / K.A. Rasinski (2000): *The Psychology of Survey Response*, Cambridge.
- Tourangeau, R. / T.W. Smith (1996): Asking Sensitive Questions: The Impact of Data Collection Mode, Question Format, and Question Context, in: *Public Opinion Quarterly* 60, S. 275-304.
- Tourangeau, R. / T. Yan (2007): Sensitive Questions in Surveys, in: *Psychological Bulletin* 133, S. 859-883.
- Traugott, M.P. / J.P. Katosh (1979): Response Validity in Surveys of Voting Behavior, in: *Public Opinion Quarterly* 43, S. 359-377.
- Van Koolwijk, J. (1969): Unangenehme Fragen. Paradigma für die Reaktionen des Befragten im Interview, in: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 21, S. 864-875.
- Voss, T. (2001): Game-Theoretical Perspectives on the Emergence of Social Norms, in: M. Hechter / K.-D. Opp (Hrsg.): *Social Norms*, New York, S. 105-136.
- Williams, R. (2007): Using heterogeneous choice models to compare logit and probit coefficients across groups, University of Notre Dame, Working Paper.
- Wyner, G.A. (1980): Response Errors in Self-Reported Number of Arrests, in: *Sociological Methods and Research* 9, S. 161-177.

Dr. Ivar Krumpal
Anatol-Fiete Näher, M.A.
Universität Leipzig
Institut für Soziologie
Beethovenstraße 15
04107 Leipzig
krumpal@sozio.uni-leipzig.de