



Max-Planck-Institut  
für ausländisches und internationales Strafrecht  
Freiburg im Breisgau

45

forschung aktuell

Franziska Kunz

**AUSWIRKUNGEN DER  
ERHEBUNGSANONYMITÄT AUF  
TEILNAHMEBEREITSCHAFT UND  
ANTWORTVERHALTEN IN  
POSTALISCHEN BEFRAGUNGEN  
ZU SELBSTBERICHTETER  
KRIMINALITÄT  
EIN METHODENEXPERIMENT**



forschung aktuell – research in brief/45

Herausgegeben von Hans-Jörg Albrecht, Albin Eser und Ulrich Sieber

**Bibliografische Information der Deutschen Nationalbibliothek**

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie;  
detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISBN 978-3-86113-172-4

Unverkäufliche Informationsbroschüre

1. Auflage 2013

Alle Rechte vorbehalten

© 2013 Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V.  
c/o Max-Planck-Institut für ausländisches und internationales Strafrecht  
Günterstalstraße 73, D - 79100 Freiburg im Breisgau Germany  
<http://www.mpicc.de>

**MAX-PLANCK-INSTITUT  
FÜR AUSLÄNDISCHES UND INTERNATIONALES  
STRAFRECHT**

**Auswirkungen der Erhebungsanonymität  
auf Teilnahmebereitschaft und  
Antwortverhalten in postalischen Befragungen  
zu selbstberichteter Kriminalität  
Ein Methodenexperiment<sup>\*</sup>**

Franziska Kunz

<sup>\*</sup> Eine frühere Version dieses Beitrages wurde am 11.09.2009 auf der 9. Jahrestagung der European Society of Criminology in Ljubljana/Slowenien vorgestellt.



# Inhaltsverzeichnis

<b>1. Einleitung</b> .....	1
<b>2. Theoretische Überlegungen und Hypothesen der Studie</b> .....	3
<b>3. Befunde empirischer Forschung</b> .....	6
<b>4. Das Methodenexperiment</b> .....	9
4.1 Grundgesamtheit und Brutto-Stichprobe .....	10
4.2 Fragebogen und Datenerhebung .....	10
4.3 Das experimentelle Design .....	11
<b>5. Ergebnisse</b> .....	12
5.1 Auswirkungen der Befragungsanonymität auf die Teilnahme- bereitschaft.....	13
5.2 Auswirkungen der Befragungsanonymität auf das Antwortverhalten .....	17
<b>6. Zusammenfassung und methodologische Empfehlungen</b> .....	22
<b>Literatur</b> .....	24



## 1. Einleitung

Die empirische Sozialforschung basiert im Wesentlichen auf Befragungsdaten und ist daher sowohl auf die Bereitschaft zur Teilnahme an Befragungen sowie auf eine aufrichtige Auskunftserteilung seitens der interessierenden Population angewiesen. Diese beiden Minimalvoraussetzungen valider Forschungsergebnisse werden zu einem wesentlichen Teil von der methodischen Vorgehensweise bei der Datenerhebung beeinflusst. Zahlreiche Aspekte spielen hier eine Rolle. Einflussgrößen und deren Wirkungsweisen zu kennen, ist daher von zentraler Bedeutung für die Forschungspraxis. Besonders relevant sind solche methodologischen Erkenntnisse für die Durchführung von Befragungen zu Themen, die seitens der Befragten als unangenehm oder heikel empfunden werden. Solche Befragungen weisen im Vergleich zu weniger „bedrohlichen“ Surveys von vorneherein ein erhöhtes Risiko von Teilnahme- bzw. Antwortverweigerungen sowie von Antwortverzerrungen im Sinne sozialer Erwünschtheit auf (Hindelang et al. 1979; Kerschke-Risch 1993, S. 21). Der vorliegende Beitrag bezieht sich auf die Erhebungsmethode der postalischen Befragung und untersucht für die Abfrage kriminalitätsbezogener Themen mithilfe eines Methodenexperiments den Einfluss der Befragungsanonymität auf die Ausschöpfungsquote<sup>1</sup> und das Antwortverhalten der Respondenten.

Geht man davon aus, dass sowohl die Teilnahme an Befragungen als auch das Antwortverhalten u.a. auf rationalen Entscheidungen seitens der angeschriebenen Personen beruhen, sollte qualitativ hochwertige Umfrageforschung so konzipiert sein, dass für die Respondenten potenzielle Nachteile minimal und Vorteile maximal sind (Dillman 1978; Esser 1986; Dillman 2000; Dillman et al. 2009). Die per-

---

<sup>1</sup> In der Praxis gilt die Ausschöpfungsquote (fälschlicherweise) weithin als zentraler Indikator für die „Repräsentativität“ der Stichprobe und damit als Gütekriterium hinsichtlich der Aussagekraft und Verallgemeinerbarkeit der erhobenen Daten (vgl. auch Schneekloth & Leven 2003, S. 19; Stoop 2005, S. 23). Dabei wird angenommen, eine höhere Ausschöpfung führe zwangsweise zu einer besseren Stichprobenqualität und steigere somit die Güte der Befragung. „Tatsächlich misst die Ausschöpfung jedoch nur [...] ,wie groß der Spielraum für Selektivität durch Nonresponse ist. Sie besagt nichts über die tatsächliche Selektivität (Deutsche Forschungsgemeinschaft 1999, S. 104)“. Nonresponse führt nur dann zu einer Verschlechterung der Erwartungstreue, wenn damit systematisch eine Unterausschöpfung von bestimmten Bevölkerungsgruppen verbunden ist“ (Schneekloth & Leven 2003, S. 19, kursiv im Original). Im Hinblick auf die für verlässliche statistische, v.a. multivariate Analysen benötigte Fallzahl bzw. Zellbesetzung ist eine höhere Ausschöpfung freilich immer besser als eine geringere. Da es zudem den Pool potenziell zu befragender Personen in der Bevölkerung (und damit auch das Image von Umfragen) schont, wenn ein benötigter Nettostichprobenumfang durch höhere Ausschöpfung einer kleineren Bruttostichprobe statt durch eine schlecht ausgeschöpfte größere Bruttostichprobe akquiriert wird, ist eine höhere Ausschöpfung auch aus forschungsethischer Sicht besser als eine geringere.

sönliche Identifizierbarkeit gegebener Antworten kann aus Sicht der Befragten mit unerwünschten Konsequenzen verbunden sein und insofern eine Bedrohung darstellen. Dies ist v.a. dann zu erwarten, wenn die abgefragten Informationen heikle Themen betreffen – die Antworten also mit gewisser Wahrscheinlichkeit sozialen und/oder rechtlichen Normen zuwiderlaufen – oder die Befragung von mächtigen Personen bzw. Institutionen durchgeführt wird, welche normabweichende bzw. unerwünschte Antworten negativ sanktionieren könnten. Zum Beispiel stellen Mitarbeiterbefragungen innerhalb von Firmen bzw. öffentlichen Institutionen (*Reips & Franek 2004*) oder die Abfrage strafbarer Handlungen in diesem Sinne „riskante“ Erhebungen dar. Von rational handelnden Befragten kann nun angenommen werden, dass sie ihre Risiken reduzieren, indem sie entweder nicht an der Befragung teilnehmen oder ihre Antworten den jeweils gültigen Normen anpassen, d.h. im Sinne sozialer Erwünschtheit antworten (*Bradburn et al. 1978*). Diese Strategien der Risikovermeidung hätten negative Auswirkungen auf die Repräsentativität der Befragung bzw. auf die Validität der Antworten.

Vollkommen anonyme Datenerhebung könnte die seitens der Respondenten vermuteten Risiken von vorneherein ausschließen, dürfte in Folge zu höherer Teilnahmebereitschaft und ehrlichen Antworten führen (*Goode & Halt 1965*, S. 172) und würde damit eine bessere Erhebungsqualität sichern. Entsprechend empfehlen zahlreiche Autoren für die Abfrage sehr sensibler, unangenehmer oder bedrohlicher Themen, die anonyme Befragungsvariante der nicht-anonymen vorzuziehen (*Singer et al. 1995*, S. 74; *Reuband 1999*, S. 59; *Dillman et al. 2009*, S. 268). Aber sind diese Überlegungen überhaupt zutreffend? Und falls ja, in welchem Ausmaß und unter welchen Bedingungen beeinflusst Anonymität die Rücklaufquote und das Antwortverhalten? Empirisch begründete, eindeutige Antworten können aus vorhandener Forschung bislang kaum abgeleitet werden. So verwundert es nicht, dass auch die Forschungspraxis – etwa im Bereich kriminologischer Erhebungen – uneinheitlich ist: Während einige Studien zu selbstberichteter Kriminalität aus forschungspraktischen Gründen unter Verwendung von Identifikationsnummern durchgeführt werden<sup>2</sup> (z.B. Schulbefragung zu Jugenddelinquenz in Köln und Freiburg; *Oberwittler et al. 2001*), gewährleisten andere Autoren die Anonymität der

---

<sup>2</sup> Diese Vorgehensweise erleichtert die Rücklaufkontrolle, vereinfacht damit die Durchführung von Nachfassaktionen und ermöglicht die Analyse von (zumindest einigen) Eigenschaften der Non-Respondenten. Durch die Identifizierbarkeit der Teilnehmer bzw. Nicht-Teilnehmer der Studie besteht ferner die Möglichkeit, eben jene im Rahmen späterer Untersuchungen (etwa Panel-Erhebungen oder Nonresponse-Studien) erneut zu kontaktieren. Gemäß entsprechender Hinweise in der Methodoliteratur (z.B. *Dillman 1978*, S. 170) wird die Verwendung von Identifikationsnummern gegenüber den Befragten offen erwähnt; zumeist wird auf deren Funktion verbesserter Rücklaufkontrolle hingewiesen. Gleichzeitig werden eine anonyme Datenauswertung und die Einhaltung von Datenschutzregelungen zugesichert, wobei dies in dezentrierter Form geschieht, um negative Folgen (*Hippler et al. 1990*) zu vermeiden.

Antworten, indem auf Identifikationsnummern verzichtet wird (z.B. allgemeine Bevölkerungsbefragung zu Werten und Kriminalität in Freiburg und Heidelberg: *Hermann* 2003) bzw. indem zusätzlich separat zurückzusendende Postkarten zur Rücklaufkontrolle eingesetzt werden (z.B. allgemeine Bevölkerungsbefragung in Bern: *Becker et al.* 2007, oder in Dresden: *Mehlkop & Becker* 2007). Die hier vorgestellte Studie prüft nun, ob und inwiefern die durch Aufdruck von Nummern erzeugte Identifizierbarkeit der Befragten deren Teilnahmebereitschaft und Antwortverhalten beeinflusst.

Zunächst werden theoretische Überlegungen zu den Auswirkungen der Befragungsanonymität auf den Rücklauf und das Antwortverhalten in postalischen Befragungen skizziert und die Hypothesen des Methodenexperiments benannt (Abschnitt 2). Abschnitt 3 gibt einen Überblick über bisherige empirische Befunde, anschließend werden die Datenbasis und das methodische Vorgehen der Studie beschrieben (Abschnitt 4). In Abschnitt 5 werden die Analyseergebnisse vorgestellt und diskutiert. Eine Zusammenfassung der Ergebnisse sowie daraus abgeleitete Empfehlungen im Hinblick auf die Konzeption postalischer Befragungen zu heiklen Themen findet der Leser im sechsten Abschnitt.

## 2. Theoretische Überlegungen und Hypothesen der Studie

Annahmen zu den Auswirkungen unterschiedlicher Modi der Erhebungsanonymität auf die Teilnahmebereitschaft und das Antwortverhalten der Respondenten lassen sich z.B. handlungstheoretisch, etwa anhand der Wert-Erwartungstheorie (*Esser* 1986) herleiten. Ausgehend von dieser Theorie werden jegliche im Rahmen einer Befragung getroffenen Handlungsentscheidungen der Respondenten – d.h. sowohl die Entscheidung zur Teilnahme an der Befragung als auch die Entscheidung für ein bestimmtes Antwortverhalten – als Ergebnis einer auf subjektiven Kosten-Nutzen-Kalkulationen beruhenden Wahl zwischen Handlungsalternativen aufgefasst. Gewählt wird vom rational handelnden Individuum letztlich die Handlungsalternative mit dem höchsten subjektiv erwarteten Nutzen bzw. den geringsten subjektiv erwarteten Kosten.

Die Beantwortung heikler bzw. bedrohlicher Fragen<sup>3</sup> kann aus Sicht der Respondenten potenziell mit Risiken<sup>4</sup>, also mit subjektiven Kosten, verbunden sein. Dies

---

<sup>3</sup> Gemeint sind damit Fragen, die sich auf sozial und/oder rechtlich genormte Sachverhalte beziehen.

<sup>4</sup> Potenzielle negative Konsequenzen stellen etwa das Erleben kognitiver Dissonanz, die Gefährdung eines positiven Selbstbildes sowie die Angst vor negativer Beurteilung durch signifikante Andere dar. Diese Handlungsfolgen können durch öffentliche Denunziation bzw. negative Sanktionierung seitens mächtiger Institutionen real werden. Letztgenanntes wiederum könnte aus Sicht der Befragten zusätzlich mit objektiven Kosten wie z.B. dem Verlust des Arbeitsplatzes oder der Auferlegung von Geldstrafe oder Freiheitsentzug verbunden sein.

dürfte v.a. dann so wahrgenommen werden, wenn die Befragten persönlich identifizierbar sind. In Bezug auf die Entscheidung für oder wider eine Befragungsteilnahme stehen den angeschriebenen Personen prinzipiell drei Handlungsalternativen offen: 1) Nicht-Teilnahme, 2) „aufrichtige“ Teilnahme und 3) „nicht-aufrichtige“ Teilnahme.

Da unter vollkommen anonymen Befragungsbedingungen die Kosten einer aufrichtigen Teilnahme im Vergleich zu nicht-anonymen Befragungen geringer sind, kann nun angenommen werden, dass bei Umfragen zu heiklen Themen anonym Befragte im Vergleich zu nicht-anonym Befragten seltener die Optionen der Nicht-Teilnahme sowie der „nicht-aufrichtigen“ Teilnahme und häufiger die Handlungsalternative „aufrichtige“ Teilnahme wählen werden. Insbesondere für die Erhebung sensibler Informationen müsste sich so insgesamt eine erhöhte Teilnahmebereitschaft unter anonymen Befragungsbedingungen ergeben. Da auch die „nicht-aufrichtige“ Teilnahme zum Rücklauf zählt und diese Handlungswahl unter den nicht-anonym Befragten häufiger auftreten dürfte, kann angenommen werden, dass die Differenz der Rücklaufquoten beider Anonymitätsmodi letztendlich nicht sehr groß ist. Die vorangegangenen Überlegungen können in folgender Hypothese festgehalten werden:

*These 1: Die Bereitschaft der Respondenten, an einer Befragung zu sensiblen Themen teilzunehmen, ist unter vollkommen anonymen Bedingungen geringfügig größer als unter nicht-anonymen Bedingungen.*

Basierend auf dem oben angeführten theoretischen Ansatz lassen sich ferner Annahmen bezüglich des Antwortverhaltens der Respondenten unter variierender Befragungsanonymität herleiten. Im Rahmen der Wert-Erwartungstheorie wird das Antwortverhalten der Befragten grundsätzlich als „[...] das kombinierte Resultat des ‚wahren Wertes‘, also der ‚personalen Identität‘ der Person auf der Zieldimension der jeweiligen Frage, und der sonstigen in der Situation aktualisierten Erwartungen und Situationsdefinitionen vor dem Hintergrund der gesamten Zielsetzungen und Interessen des Befragten“ (Esser 1986, S. 320) aufgefasst. In einer konkreten Fragesituation verfügen Respondenten nun über folgende vier Handlungsalternativen: 1) Antwortverweigerung, 2) Angabe des „wahren Wertes“, 3) Angabe eines „unwahren“ Wertes (kann sozial erwünschter Wert sein) sowie 4) Angabe eines zufälligen Wertes. Sozial erwünschte Antwortverzerrungen im Sinne der bewussten Abweichung von einem „wahren Wert“ können dabei einerseits aufgrund eines in der Persönlichkeit verankerten Bedürfnisses nach sozialer Anerkennung („need for social approval“) und andererseits infolge von Umständen der Datenerhebung, die negative Konsequenzen befürchten lassen, erfolgen. Anders ausgedrückt: „Die motivationale Basis sozial erwünschter Reaktionen besteht in einem generellen Bedürfnis nach sozialer Anerkennung und/oder (!) in der *sonstwie* erlebten Relevanz

der Situation für die Interessen des Befragten, wie sie durch die Situation [...] signalisiert wird“ (Esser 1986, S. 326; kursiv im Original). Für beide „Motivations-Typen“ gilt, dass die Erwünschtheitsneigung erst durch die (potenzielle) Öffentlichkeit der Situation „aktiviert“ wird (vgl. Esser 1986, S. 326) – schließlich kann eine anonyme Antwort weder positive (z.B. soziale Anerkennung) noch negative Handlungskonsequenzen für das Subjekt beinhalten. Das Ausmaß sozial erwünschten Antwortverhaltens der Befragten bei heiklen Themen hängt also „[...] v.a. von der (vermuteten) Überprüfbarkeit der Angaben und von der (vermuteten) Vertraulichkeit der Situation und Datenbehandlung ab“ (Esser 1986, S. 318; runde Klammern im Original).

Unter identifizierbaren Erhebungsbedingungen könnte sich daher für einige Respondenten eine vom „wahren Wert“ abweichende Antwort, d.h. eine der Antwortoptionen 1, 3 bzw. 4 (siehe oben), als optimale Handlungsalternative ergeben. Bei anonymer Befragung erscheinen diese Handlungsoptionen irrational. Empirisch sollen entsprechend folgende Thesen geprüft werden:

*These 2: Das Ausmaß von Antwortverweigerungen ist in Bezug auf heikle Fragen unter anonymen Bedingungen geringer als unter nicht-anonymen Bedingungen.*

*These 3: Die Bereitschaft der Respondenten, auf heikle Fragen „aufrichtig“, d.h. mit dem „wahren Wert“ zu antworten, ist unter anonymen Bedingungen größer als unter nicht-anonymen Bedingungen.*

Wie bereits weiter oben festgelegt, soll hier eine Frage als heikel gelten, wenn deren Inhalt sozial und/oder rechtlich genormt ist. Sofern die Antworten auf so definierte Fragen in der anonymen Befragungsgruppe im Vergleich zur nicht-anonymen weniger sozial erwünscht sind, also sozialen/rechtlichen Normen weniger entsprechen, wird These 3 beibehalten.<sup>5</sup>

---

<sup>5</sup> Nun kann der „wahre Wert“ sowohl sozial erwünscht als auch sozial unerwünscht sein (Esser 1986, S. 327 ff.). Sofern z.B. die Befragten im Hinblick auf die untersuchten „heiklen Fragen“ nichts zu verbergen hätten, d.h. sofern ihre sozial erwünschte Antwort zugleich eine aufrichtige ist, wäre die Hypothese nicht prüfbar. Die These ist nur unter der Annahme prüfbar, dass die abgefragten Einstellungen bzw. Verhaltensweisen zumindest bei einem Teil der Respondenten in sozial unerwünschter Weise auftreten bzw. aufgetreten sind. Aufgrund der Vielzahl heikler Fragen, anhand derer die These geprüft werden soll, dürfte diese Annahme m.E. mit einer hinreichend hohen Wahrscheinlichkeit zutreffen.

### 3. Befunde empirischer Forschung

Empirische Forschung zu den Auswirkungen anonymer Befragung erbrachte bislang uneinheitliche Befunde. Bezüglich des Einflusses von Anonymität auf die *Teilnahmebereitschaft* ergibt sich folgendes Bild: Einige Studien stellen eine *erhöhte Teilnahmerate* unter anonymen Bedingungen fest (King 1970; Jones 1979; Futrell 1981; Singer et al. 1995), die Mehrzahl der Untersuchungen findet allerdings *keine Unterschiede* (Ash & Abramson 1952; Fuller 1974; Kanuk & Berenson 1975; Futrell & Swan 1977; Wildman 1977; Yu & Cooper 1983; Albaum 1987; Wolfgang et al. 1987; Campbell & Waters 1990; Yammarino et al. 1991; Reuband 1999; Blumberg et al. 2004; Reuband 2006). Die Studie von Fuller (1974) erbringt gemischte, auf je spezifische Subpopulationen bezogene Ergebnisse. Für Mitarbeiterbefragungen weisen Reips & Franek (2004) eine Tendenz zu *geringerer Rücklaufquote* unter vollständiger Anonymität nach. Hippler et al. (1990) berichten ferner, dass ausführliche Vertraulichkeitshinweise bzw. die Betonung der Befragungsanonymität Teilnahmequoten senken, sofern nicht-sensitive Themen abgefragt werden. Die beiden letztgenannten Befunde sind zwar im Rahmen der Anonymitätsforschung relevant. Auf die hier untersuchte Befragungssituation dürften diese Ergebnisse allerdings kaum übertragbar sein, da es sich bei vorliegender Befragung nicht um eine Mitarbeiterbefragung handelt<sup>6</sup> und das Schwerpunktthema „Kriminalität“ zudem als heikel gelten kann.

Ähnlich widersprüchlich sind die Forschungsergebnisse im Hinblick auf das *Antwortverhalten*. Zwar finden sich hier überwiegend empirische Belege für die Annahme, dass (bei heiklen Themen) sozial erwünschtes Antwortverhalten durch vollständige Anonymität reduziert wird (Olson 1936; Benson 1941; Fischer 1946; Becker & Bakal 1970; King 1970; Tyagi 1989; Reuband 1999; Ong & Weiss 2000; Reips & Franek 2004; van de Looij-Jansen et al. 2006). Daneben gibt es aber auch Studien, die *keine Unterschiede* finden (Ash & Abramson 1952; Berman et al. 1977; Reuband 2011) oder *theoriekonträre Resultate* erbringen, wie etwa eine Untersuchung von Epperson & Peck (1977), in der die Teilnehmer eines Pflichtprogramms zur Verkehrssicherheit diese vom Department of Motor Vehicles (DMV) durchgeführte Maßnahme unter anonymen Befragungsbedingungen signifikant positiver bewerten als unter nicht-anonymen Bedingungen.

Die Mehrzahl der erwähnten Methodenexperimente wurden in den 1970-/1980er Jahren in den USA durchgeführt. Eine Übertragbarkeit dieser Befunde auf heutige,

---

<sup>6</sup> Es ist zu vermuten, dass bei Mitarbeiterbefragungen die Rücklaufquoten unter anonymen Bedingungen u.a. deshalb geringer sind als unter nicht-anonymen Bedingungen, weil die Mitarbeiter die Befragung als zusätzliche und evtl. unbezahlte Arbeitsleistung für ihren Arbeitgeber ansehen. Da bei anonymer Befragung nicht überprüft werden kann, wer an der Befragung teilnimmt, ist der Anreiz zur Verweigerung somit höher. Diese Kosten-Nutzen-Interpretation der Befragten dürfte sich nicht ohne weiteres auf allgemeine Bevölkerungsbefragungen übertragen lassen.

zudem europäische Verhältnisse ist fragwürdig. Nur wenige Untersuchungen (*Reuband* 1999; *Reips & Franek* 2004; *Reuband* 2006; 2011) beziehen sich auf den deutschsprachigen Raum. *Reuband* (1999; 2006; 2011) liefert die aktuellsten Befunde für Deutschland, die Untersuchung von 2006 bezieht sich zudem als einzige speziell auf ältere Menschen, was bedeutsam für das vorzustellende Methodenexperiment ist. Die Ergebnisse der *Reuband*-Studien werden deshalb kurz gesondert vorgestellt:

In den Untersuchungen von 1999 und 2006 zeigte sich kein Effekt der Anonymität auf die Response-Rate. Was das Antwortverhalten der Befragten bei sensiblen Themen betrifft, gelangt *Reuband* in zwei Untersuchungen<sup>7</sup> (1999; 2011) zu unterschiedlichen Resultaten. In der ersten Untersuchung, die in der westdeutschen Großstadt Düsseldorf und in der ostdeutschen Großstadt Dresden durchgeführt wurde (*Reuband* 1999), ist für Düsseldorf (nicht aber für Dresden) ein Anonymitätseffekt zu beobachten: „[...] hier [...] scheint die anonyme Befragungsvariante – wenn auch in der Regel schwache – Auswirkungen auszuüben und offenere Antworten zu begünstigen“ (*Reuband* 1999, S. 59).

In der zweiten Untersuchung, einer komplett anonymen Mehrthemenbefragung, untersuchte *Reuband* (2011) in fünf deutschen Großstädten (München, Hamburg, Kiel, Stuttgart und Dresden; n = 4.831), in welchem Ausmaß die Befragten zur Preisgabe ihrer Kennziffer und damit zur freiwilligen De-Anonymisierung ihrer Antworten bereit sind und inwiefern sich de-anonymisierungsbereite von hierzu nicht-bereiten Personen hinsichtlich verschiedener Merkmale unterscheiden. Als potentielle Unterscheidungsmerkmale wurden soziodemographische Variablen (Geschlecht, Alter, Bildung), das Politikinteresse sowie – für den vorliegenden Beitrag besonders relevant – einige Angaben zur eigenen Delinquenz analysiert. Erfasst wurden hier insgesamt fünf Sachverhalte: Für die Delikte Schwarzfahren, Ladendiebstahl und Drogengebrauch wurde jeweils gefragt, ob die Befragten dies schon *jamals in ihrem Leben* getan haben, darüber hinaus wurde der ungefähre Zeitpunkt des letzten Drogenkonsums sowie die Bereitschaft, zukünftig Drogen zu nehmen (wenn diese auf einer Party angeboten würden), erhoben. Insgesamt stellt *Reuband* (2011, S. 30) anhand der vorliegenden Daten keine nennenswerten Unterschiede zwischen de-anonymisierten und anonymen Befragten fest – „weder hinsichtlich der sozialen Zusammensetzung noch des Eingeständnisses der selbst verübten Delinquenz und der Delinquenzbereitschaft“. Allerdings dürfte die Aussagekraft dieser Analysen im Hinblick auf die Frage, ob Anonymität das Antwortverhalten bei sensiblen Fragen beeinflusst, insgesamt recht begrenzt sein. *Reuband* (2011, S. 30) selbst wendet diesbezüglich ein, „[...]“, dass die erfragten Delikte relativ trivial, in der Jugendphase relativ weit verbreitet sind [...] und es im Erwachsenenalter – subjektiv gesehen – keinerlei nennenswerte Bedrohung mehr darstellt, wenn man deviante Verhaltens-

---

<sup>7</sup> Die *Reuband*-Studie von 2006, in der Senioren zur NS-Zeit befragt wurden, beinhaltet keine Analysen dazu, inwiefern das Antwortverhalten älterer Menschen vom Grad der Anonymität beeinflusst wird.

weisen aus der Jugendzeit eingesteht. Im Fall des Drogengebrauchs hat sich das gesellschaftliche Klima zudem grundlegend gewandelt, Cannabiskonsum gilt nicht mehr als so schlimm wie vor wenigen Jahrzehnten [...]. Sich als ehemaliger oder aktueller Konsument zu erkennen zu geben, ruft in der Gesellschaft nur noch bedingt heftige Reaktionen hervor.“

Den vorgestellten Daten lässt sich jedenfalls entnehmen, dass die Beantwortung der Frage nach dem Zeitpunkt des letzten Drogenkonsums – die wohl eine der (wenigen) Fragen darstellt, die von den Befragten noch am ehesten als heikel empfunden wurden – in München und Dresden durchaus unterschiedlich ausfällt, wobei die anonym Gebliebenen gegenüber den de-anonymisierten Personen jeweils häufiger berichten, innerhalb des letzten Jahres bzw. innerhalb der letzten 30 Tage Drogen konsumiert zu haben<sup>8</sup> (*Reuband* 2011, S. 27, Tab. 2 sowie S. 28, Fn. 19). Zwar erwähnt *Reuband* dies selbst nicht, seine Ergebnisse zeigen aber ferner, dass – zumindest in Dresden – auch die Bereitschaft zum zukünftigen Drogengebrauch in Abhängigkeit der Befragungsanonymität variiert: Unter den anonym Befragten befinden sich (relativ gesehen) deutlich mehr potentielle Drogenkonsumenten als unter den Befragten, die ihre Antworten de-anonymisiert hatten.<sup>9</sup>

Die einschlägige Methodenforschung weist insgesamt eine enorme Vielfalt hinsichtlich der methodischen Konzeptionen, der zugrunde liegenden Populationen, der Stichproben und der untersuchten Befragungsthemen auf. Lediglich drei der benannten Studien analysieren die Auswirkungen von Anonymität in postalischen Befragungen zu Themen, die gemeinhin als heikel gelten und damit dem hier untersuchten Thema „Delinquenz“ noch am ähnlichsten sind bzw. dieses sogar berücksichtigen:

*King* (1970) führte eine postalische Befragung zum Drogenkonsum von Jugendlichen durch; *Ong & Weiss* (2000) untersuchten Effekte von Anonymität an einer Stichprobe von Studenten u.a. zu den Themen Betrug in Prüfungen, Masturbation und Ladendiebstahl. *Van de Looij-Jansen et al.* (2006) bezogen ihr Experiment auf das Gesundheitsverhalten von Jugendlichen, wobei auch Alkoholkonsum und aggressives Verhalten thematisiert wurden. Alle Studien fanden weniger sozial erwünschte Antworten unter anonymen Bedingungen. Am klarsten sind die Effekte bei *Ong & Weiss* (2000, S. 1701): „The results show that anonymity and confidentiality should not be seen as interchangeable. Anonymity induced many more revelations.“ *King* (1970), der als einziger die Rücklaufquote in Abhängigkeit der Anonymität auswertete, fand in zwei Gruppen, die zu Marihuana- bzw. LSD-Gebrauch Auskunft geben sollten, je erhöhte Rücklaufquoten in der anonymen Kondition. Alle

---

<sup>8</sup> In Hamburg, Kiel und Stuttgart zeigt sich dieser Effekt nicht.

<sup>9</sup> In Dresden geben 39 % der anonym Befragten an, „auf jeden Fall/sehr oder eher wahrscheinlich“ auf einer Party Haschisch zu nehmen, wenn es ihnen ein Bekannter anbieten würde, im Vergleich zu 22 % der nicht-anonym Befragten (*Reuband* 2011, S. 27, Tab. 2).

Respondenten zusammen betrachtet, ergab sich hier unter anonymen Bedingungen eine um 4 % höhere Ausschöpfungsquote.<sup>10</sup> Zu den Effekten der Anonymität in postalischen Befragungen speziell zu selbstberichteter Kriminalität liegen außer den Ergebnissen von *Ong & Weiss* (2000) zum Ladendiebstahl (gemessen mit einem Item) bislang keine weiteren empirischen Erkenntnisse vor.

In Anbetracht der Vielfältigkeit relevanter Forschung und Forschungsbefunde wird v.a. eines offensichtlich: Auswirkungen von Anonymität stellen ein komplexes Phänomen dar. Zahlreiche Faktoren scheinen die Bedeutung anonymer im Vergleich zu nicht-anonymer Erhebung zu moderieren, genauere Kenntnisse hierzu liegen allerdings kaum vor. Zumindest fand *Jones* (1979) in den Vereinigten Staaten Hinweise darauf, dass Anonymität in Populationen mit höherem Bildungsstand und Einkommen zu einer gesteigerten Ausschöpfungsquote führt, während der gegenteilige Effekt in Wohngebieten festgestellt wurde, die durch eine hohe Zu- und Wegzugsrate gekennzeichnet sind.

Insgesamt deuten die Forschungsergebnisse darauf hin, dass die Ausschöpfungsquote bei wenig heiklen Befragungsthemen nicht und bei eher sensiblen Themen schwach positiv von der Befragungsanonymität beeinflusst wird. Inwiefern das Antwortverhalten der Respondenten mit Anonymitätsgewährleistungen zusammenhängt, scheint ebenfalls stark vom Befragungsthema bzw. der Sensitivität der abgefragten Informationen abzuhängen. Die wenigen Studien zu normabweichendem Verhalten erbrachten zumindest teilweise Belege dafür, dass Antworten auf heikle Fragen „offener“, d.h. weniger sozial erwünscht sind, wenn die Befragung anonym ist.

#### 4. Das Methodenexperiment

Das Methodenexperiment, von dem hier berichtet wird, fand im Rahmen einer regionalen postalischen Befragung zu selbstberichteter Kriminalität im höheren Lebensalter statt, die im Jahr 2009 am Freiburger Max-Planck-Institut für ausländisches und internationales Strafrecht in Kooperation mit dem Institut für Soziologie der Universität Freiburg durchgeführt wurde.<sup>11</sup> Um das Erhebungsinstrument der postalischen Befragung zu optimieren und potenzielle Probleme der Datenerhebung

---

<sup>10</sup> Dieses Ergebnis ist nicht signifikant. Allerdings ist die Teststärke der Studie für gesicherte Aussagen zur Nicht-Signifikanz dieses Effekts zu gering. Zahlreiche Methodenexperimente – auch das hier vorgestellte, siehe Abschnitt 5.1 – weisen im Hinblick auf einzelne Effekte zu geringe Teststärken auf (ausführlicher zur statistischen Bedeutung und zur Berechnung der Teststärke z.B. *Cohen* 1977; *Cohen et al.* 2003, S. 15, 652 ff.).

<sup>11</sup> Für eine umfassende methodologische und inhaltliche Vorstellung dieser Befragung siehe *Kunz* 2013b; knappe Darstellungen wesentlicher Befunde finden sich bei *Kunz* 2011; 2013a.

zu identifizieren, wurde im Vorfeld ein „Zwei-Phasen-Pretesting“ (*Priifer & Rex-roth* 2000) durchgeführt. In der ersten Phase wurde der Fragebogen mittels kognitiver Pretests in Form mündlicher Interviews getestet. Mit dem modifizierten Fragebogen wurde in der zweiten Phase ein sog. Standard-Pretest, d.h. die Prüfung des Erhebungsinstruments unter den für die Hauptidehebung geplanten Konditionen durchgeführt. Dieser Standard-Pretest beinhaltet das Methodenexperiment, das im vorliegenden Beitrag vorgestellt wird.

#### **4.1 Grundgesamtheit und Brutto-Stichprobe**

Die Grundgesamtheit des Standard-Pretests bildeten deutsche Staatsbürger, die zwischen dem 01.01.1929 und dem 01.01.1959 geboren wurden, mit privatem Wohnsitz im Regierungsbezirk Freiburg, d.h. in der Region Südbaden.<sup>12</sup> Die Brutto-Stichprobe der postalischen Befragung beruhte auf einer geschichteten, geklumpten, zweistufigen Zufallsauswahl und umfasste 400 Personen aus 12 zufällig ausgewählten Gemeinden des Regierungsbezirkes Freiburg.

#### **4.2 Fragebogen und Datenerhebung**

Der dem Standard-Pretest zugrunde liegende Fragebogen umfasst 24 Seiten mit insgesamt 85 Fragen bzw. 240 einzelnen Items.<sup>13</sup> Die Befragung trug den Titel „Alltag & Gesetze. Erfahrungen und Einstellungen in der zweiten Lebenshälfte“ und beinhaltete u.a. Fragen zu Einstellungen gegenüber Gesetzen und Gesetzesübertretungen, zu Viktimisierungserfahrungen, zu Kriminalitätsfurcht sowie zu bisheriger polizeilicher Registrierung, kriminellen Verhalten innerhalb verschiedener Lebensabschnitte und zu delikt-spezifischer Delinquenzbereitschaft. Die Datenerhebung fand im Mai/Juni 2009 statt und erstreckte sich über einen Zeitraum von vier Wochen. Die Befragten erhielten zunächst ein doppelseitiges Anschreiben, den Fragebogen, einen adressierten und frankierten Rückumschlag sowie eine adressierte und frankierte Antwortpostkarte.<sup>14</sup> Zur Steigerung des Rücklaufs wurde allen Angeschriebenen „als kleine Aufwandsentschädigung und Dankeschön“ be-

---

<sup>12</sup> Für Menschen, die in Institutionen wie etwa Alters- oder Pflegeheimen leben, sowie für Personen ohne deutsche Staatsbürgerschaft können im Rahmen der Studie entsprechend keine Aussagen getroffen werden.

<sup>13</sup> Davon sollten 76 Fragen bzw. 164 Items von allen Befragten beantwortet werden, während weitere neun Fragen bzw. 76 Items als sog. Filterfragen nur in Abhängigkeit vorheriger Antworten ggf. zu beantworten waren.

<sup>14</sup> Diese Antwortpostkarte diente einerseits der Rücklaufkontrolle, andererseits wurde hier die Bereitschaft der Respondenten zur Teilnahme an einem Interview im Rahmen einer Anschlussstudie zum Thema „Moral“ erfragt (zu dieser Studie siehe *Kunz* 2014 [im Erscheinen]).

dingungslos und im Voraus eine Fünf-Euro-Banknote beigelegt.<sup>15</sup> Diese Gabe konnte in jedem Fall behalten werden. Zwei Wochen nach dem Erstversand wurde ein kurzes Schreiben an alle Respondenten verschickt. Dieses enthielt sowohl die erneute Bitte um Teilnahme als auch – für den Fall der bereits erfolgten Rücksendung des Fragebogens – Dank für die geleistete Unterstützung.

### 4.3 Das experimentelle Design

Um die Effekte der Erhebungsanonymität auf den Rücklauf und das Antwortverhalten zu untersuchen, wurden zwei Variationen geprüft:

- 1) anonyme Befragung,
- 2) nicht-anonyme Befragung.

Beide Befragungsvarianten unterschieden sich in Bezug auf zwei Aspekte voneinander:

- 1) Der Fragebogen der *nicht-anonymen* Befragung enthielt eine fünfstellige Zahl, welche mit den personenbezogenen Daten der Befragten verknüpft war. Diese Identifikationsnummer war auf der Innenseite der hinteren Umschlaghülle des Fragebogens am unteren Ende aufgedruckt. Die *anonyme* Befragungsvariante enthielt *keine* Nummer.
- 2) Im Anschreiben der *nicht-anonymen* Befragung wurden Anonymität und Vertraulichkeit der Angaben wie folgt thematisiert: „Wir wollen nur wissen, wie Menschen in der 2. Lebenshälfte über unser Thema denken und nicht, wer genau was denkt. Die Nummer, die am Ende des Fragebogens unten aufgedruckt ist, dient lediglich der besseren Übersicht über den Rücklauf der Befragung. **Ihre Angaben werden von uns vollkommen anonym ausgewertet und nicht weitergegeben.**“ Im Anschreiben zur *anonymen* Befragung hieß es stattdessen: „**Auf dem Fragebogen befindet sich nirgendwo Ihr Name oder Ihre Anschrift!** Wir wollen nur wissen, wie Menschen in der 2. Lebenshälfte über unser Thema denken und nicht, wer genau was denkt. Es ist für uns nicht erkennbar, wer welchen Fragebogen ausgefüllt hat. **Ihre Angaben werden also vollkommen anonym ausgewertet und nicht weitergegeben.**“ (Fettdruck in den originalen Anschreiben).

---

<sup>15</sup> Der Einsatz von Anreizen bei der Durchführung postalischer Befragungen wird u.a. von Dillman (2000, S. 150) und Dillman *et al.* (2009, S. 238 ff.) im Rahmen der Tailored Design Method empfohlen. Für empirische Forschungsbefunde und weitere Literaturhinweise zur *allgemeinen* Wirkung von Incentives siehe Arzheimer & Klein 1998; Singer 1998; Singer *et al.* 1998; Diekmann & Jann 2001; Porst 1999; Stadtmüller & Porst 2005 und Stadtmüller 2009. Speziell zur Wirkung vorausbezahlter monetärer Anreize auf die Response-Raten, auf soziodemographische Respondentenmerkmale sowie auf die selbstberichtete Delinquenz in postalischen *Kriminalitätsbefragungen* siehe die empirischen Befunde von Becker *et al.* 2007; Mehlkop & Becker 2007 sowie Becker & Mehlkop 2011.

Um bei allen Survey-Respondenten die Bereitschaft zur Teilnahme an einem Interview der geplanten Anschlussstudie (zum Thema „Moral“) zu erfragen, lag die adressierte und frankierte Rückantwortkarte, welche jene Frage enthielt, *beiden* Befragungsvarianten bei. Im Methodenexperiment wurde neben den Auswirkungen der Anonymität zudem ein Fragereihenfolge-Effekt untersucht.<sup>16</sup> Es ergaben sich somit vier experimentelle Variationen, d.h. ein 2x2-faktorielles Design. Von den 400 anzuschreibenden Personen wurden je 100 zufällig ausgewählt und den experimentellen Bedingungen zugeordnet, so dass vier Subsamples entstanden. Der vorliegende Beitrag berichtet lediglich die Ergebnisse für die Variation der Befragungsanonymität. Je 200 Personen wurden anonym bzw. nicht-anonym befragt.

## 5. Ergebnisse

Der Standard-Pretest erzielte eine Netto-Rücklaufquote in Höhe von 49,9 % (Tabelle 1). Während 196 Fragebögen eingingen, wurden lediglich 148 Postkarten zurückgesandt.<sup>17</sup> Unter der Annahme, die Antwortkarten wurden nur dann zurückgeschickt, wenn auch ein Fragebogen ausgefüllt wurde, entspricht dies einer Benachrichtigungsquote von 75,5 %.<sup>18</sup>

*Tabelle 1: Rücklauf und Ausfälle der postalischen Befragung des Standard-Pretest*

Brutto-Stichprobe (N)	400
neutrale Ausfälle*	7
bereinigte Brutto-Stichprobe (ber. N)	393
Nicht-Teilnahme**	197
auswertbare Fragebögen (n)	196
<b>Netto-Ausschöpfungsquote</b>	<b>49,9 %</b>

\* Als „neutrale Ausfälle“ gelten hier: unbekannt/verzogen, verstorben, Krankheit/Pflege/Behinderung, unzureichende deutsche Sprachkenntnisse, Person gehört nicht zur Grundgesamtheit.

\*\* Unter „Nicht-Teilnahme“ zählen: keine Reaktion, explizite Verweigerungen und unausgefüllt zurückgeschickte Fragebögen.

<sup>16</sup> Es wurde exemplarisch für die beiden Delikte Ladendiebstahl und Versicherungsbruch untersucht, inwiefern in (postalischen) Befragungen ein variierender Fragekontext die Antworten der Respondenten auf Fragen zur Norminternalisierung und Delinquenzbereitschaft beeinflusst; ausführlicher dazu siehe *Kunz & Grundies* 2011.

<sup>17</sup> Die Aussage im Lehrbuch von *Schnell et al.* (2008, S. 363), wonach die Postkartentechnik wenig erfolgversprechend sei, weil „[v]iele Befragte [...] zwar die Postkarte, nicht jedoch den Fragebogen zurück[schicken]“, bestätigt sich damit in der vorliegenden Untersuchung (wie auch in anderen Studien, z.B. *Reuband* 1999; *Kunz* 2013 [im Erscheinen]) nicht.

<sup>18</sup> Diese Quote liegt deutlich über bisher dokumentierten Erfahrungen, etwa den von *Reuband* (1999, S. 58) berichteten Raten für allgemeine Bevölkerungsbefragungen in Dresden (66 %) und Düsseldorf (60 %) und ist angesichts der älteren Population und der zusätzlichen Komplexität der Aufgabenstellung sehr zufriedenstellend.

Im Folgenden werden die Resultate des Methodenexperiments in Bezug auf die Teilnahmebereitschaft (Abschnitt 5.1) und das Antwortverhalten der Respondenten (Abschnitt 5.2) vorgestellt.

## 5.1 Auswirkungen der Befragungsanonymität auf die Teilnahmebereitschaft

Tabelle 2 zeigt, dass die Ausschöpfungsquote unter anonymen Bedingungen mit 52,3 % höher ist als unter nicht-anonymen Bedingungen (47,5 %). Die Differenz der Response-Raten beträgt 4,8 % und ist damit wie erwartet relativ gering. Zwar ist der beobachtete Unterschied hier nicht signifikant.<sup>19</sup> Da die Teststärke der durchgeführten Analyse aufgrund des geringen Stichprobenumfangs jedoch zu gering ist, um Unterschiede der beobachteten Größenordnung verlässlich als insignifikant auszuweisen,<sup>20</sup> wird Hypothese 1 hier trotz des negativen Testresultats, das üblicherweise ein Ablehnungskriterium darstellt, beibehalten.

*Tabelle 2: Teilnahmebereitschaft in Abhängigkeit der Befragungsanonymität (in Klammern: Zeilenprozente\*)*

	Teilnahme	Verweigerung	Gesamt
Anonymität versprochen	94 (47,5 %)	104 (52,5 %)	198 (100,0 %)
Anonymität gewährleistet	102 (52,3 %)	93 (47,7 %)	195 (100,0 %)
Gesamt	196 (49,9 %)	197 (50,1 %)	393 (100,0 %)

\* Die Prozentangaben beziehen sich jeweils auf die bereinigte Brutto-Stichprobe, d.h. neutrale Ausfälle wurden zuvor ausgeschlossen.

Ein Abgleich mit den Befunden von *King* (1970), der bei einer Befragung Jugendlicher zu deren Drogenkonsum eine um 4 % erhöhte Rücklaufquote unter anonymen Bedingungen beobachtete, kann als Indiz dafür angesehen werden, dass das hier berichtete Ergebnis kein zufälliges ist, sondern dass ein schwacher aber positiver Zusammenhang zwischen Anonymität und Teilnahmebereitschaft für heikle Befragungen in der Grundgesamtheit tatsächlich besteht.

<sup>19</sup>  $\Phi = -0,048$ ;  $p_{2\text{-seitig}} = 0,338$ ;  $df = 1$ ,  $n = 393$ .

<sup>20</sup> Die durchgeführte Analyse (mit  $N = 393$ ) weist für die beobachtete Effektstärke (Cohen's  $d = 0,097$ ; entspricht 4,8 % Rücklaufdifferenz) lediglich eine Teststärke in Höhe von 0,158 für zweiseitige Tests bzw. in Höhe von 0,247 für einseitige Tests auf. Das gemeinhin übliche Niveau von 0,8 wird damit weit verfehlt. Da somit der Beta-Fehler, d.h. die Wahrscheinlichkeit, mit der die Nullhypothese fälschlicherweise beibehalten wird, entsprechend stark überhöht ist, liefert ein Signifikanztest keine verlässlichen Negativ-Ergebnisse.

*Wie kann der schwache Einfluss von Anonymität auf die Teilnahmebereitschaft begründet werden?*

Eine mögliche Erklärung für die geringe Differenz der Teilnahmequoten bietet die auf der Wert-Erwartungstheorie basierende Vermutung, dass unter nicht-anonymen Bedingungen der Anteil „nicht-aufrichtiger“ Teilnahme höher ist als unter anonymen Bedingungen, sich die unterschiedlichen Anteile „aufrichtiger“ Teilnahme entsprechend nur relativ schwach auf die gesamte Teilnahmequote auswirken (vgl. die Argumentation unter Abschnitt 2).

Es könnte jedoch auch sein, dass der beobachtete Effekt deshalb gering ist, weil die postalische Befragung aufgrund ihrer Administrationsform von vornherein eine recht hohe Anonymität des Befragten gewährleistet. Sie gilt ja gerade aus diesem Grund als besonders geeignet für die Erhebung sensibler Informationen (Baurmann et al. 1991, S. 161; Kury 1994; Schnell et al. 1995, S. 333; Reuband 2000; Tourangeau et al. 2000, S. 205 ff.; Reuband 2002; Becker & Günther 2004, S. 56).

Eine weitere Vermutung ist die, dass zwischen der Vergabe von Incentives und der Befragungsanonymität eine negative Interaktion besteht, dass also monetäre Anreize die Kooperationsbereitschaft derart positiv beeinflussen, dass die Bedeutung von Anonymität reduziert wird. Dass Befragte durch die Beigabe eines Geldgeschenkes dazu motiviert werden können, den Fragebogen ausführlicher, weniger fehlerhaft und vollständiger auszufüllen (Wotruba 1966; McDaniel & Rao 1980; James & Bolstein 1990) sowie zügiger zu retournieren (Brennan 1992; Arzheimer & Klein 1998), ist bekannt. Bezugnehmend auf die soziale Austauschtheorie (Homans 1961; Blau 1964), die auch Dillman (1978) zur Erklärung der Teilnahme an postalischen Befragungen heranzieht, könnte analog vermutet werden, dass fehlende Anonymität der Antworten eher in Kauf genommen wird, wenn die Teilnahme durch ein Geldgeschenk vorab honoriert wird. Dies wiederum kann wie folgt erklärt werden: Die Vorleistung des Incentives, welcher als Vertrauensvorschuss interpretiert wird, evoziert beim Empfänger das Gefühl der Verpflichtung, sich für das Geschenk erkenntlich zu zeigen. Die Bestrebung seitens der Respondenten, gemäß der Reziprozitätsnorm (Gouldner 1960; Diekmann & Jann 2001) eine adäquate Gegenleistung für die erhaltene Gabe zu erbringen (Mauss 1985 [1925]), könnte nun sowohl die Teilnahme an der Befragung als auch die „aufrichtige“ Beantwortung der gestellten (heiklen) Fragen umfassen<sup>21</sup> – und zwar weitgehend unabhängig von der Erhebungsanonymität.

Leider können die angeführten Überlegungen im Rahmen dieser Studie nicht näher untersucht werden. Lediglich die von Bradburn et al. (1978) geäußerte These, wonach die Befragten unabhängig von der Anonymität der Befragung zwar teilnehmen, ihre Antworten jedoch der Identifizierbarkeit anpassen, kann hier geprüft werden. Dafür wird zunächst anhand einiger soziodemographischer Variablen

---

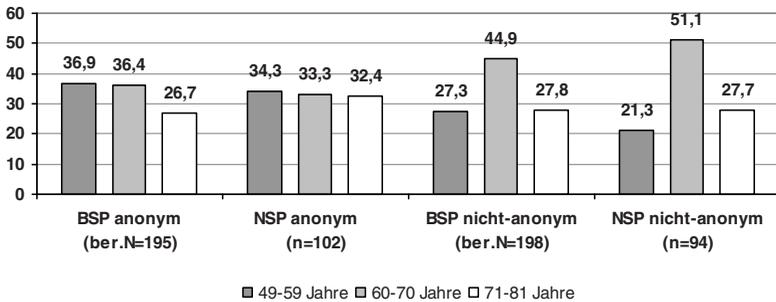
<sup>21</sup> Eine ähnliche Annahme findet sich auch bei Mehlkop & Becker 2007, S. 6.

überprüft, ob sich die Respondenten der anonymen und nicht-anonymen Befragungsvariante systematisch voneinander unterscheiden. Eine solche Analyse gibt Aufschluss darüber, inwiefern sich die Erhebungsanonymität auf die Zusammensetzung der Befragtenstichprobe auswirkt. In einem zweiten Schritt wird dann das inhaltliche Antwortverhalten der Respondenten in Abhängigkeit der Anonymität ihrer Antworten näher untersucht (Abschnitt 5.2).

Hinsichtlich der Merkmale Bildung, Einkommen, soziale Selbsteinstufung, Wohngegend und Familienstand konnten keine signifikanten Unterschiede zwischen anonymer und nicht-anonymer Netto-Stichprobe festgestellt werden. Allerdings sind die Teilnehmer der nicht-anonymen Kondition im Durchschnitt älter (Mean = 66,0 Jahre) als diejenigen der anonymen Kondition (Mean = 64,4 Jahre).<sup>22</sup>

In Abbildung 1 sind die Altersverteilungen in Brutto- und Netto-Stichprobe der anonymen und nicht-anonymen Befragungsvariante dargestellt.

Abbildung 1: Altersverteilung (gruppiert) in der Brutto-Stichprobe (BSP) und Netto-Stichprobe (NSP) nach Befragungsanonymität (in Prozent)



Zwar unterscheiden sich bereits die Brutto-Stichproben beider Befragungsvarianten hinsichtlich der Altersverteilung etwas voneinander. Die jeweiligen Netto-Stichproben zeigen aber deutlich größere Unterschiede. In der anonymen Befragung konnten Teilnehmer der drei Altersgruppen annähernd proportional zur Ausgangsstichprobe rekrutiert werden. Die Netto-Stichprobe der nicht-anonymen Befragung hingegen wird im Vergleich zur Brutto-Stichprobe stark überproportional von Teilnehmern der mittleren Alterskategorie dominiert, während die 49- bis 59-Jährigen unterrepräsentiert sind. Für ein besseres Verständnis dieser Verteilungen sind in Tabelle 3 ergänzend die altersgruppenbezogenen Response-Raten der anonymen und nicht-anonymen Befragung aufgeführt.

<sup>22</sup> Für die gruppierte Altersvariable ist der Unterschied zwischen den Teilnehmern der anonymen und nicht-anonymen Befragung signifikant (Chi-Quadrat = 6,997; Cramer's V = 0,19; df = 2; p<sub>2-seitig</sub> = 0,03; n = 196).

*Tabelle 3: Netto-Response-Raten nach Altersgruppen (AG) und Befragungsanonymität (in Klammern: absolute Anzahl NSP/BSP)*

	AG1 49-59 Jahre	AG2 60-70 Jahre	AG3 71-81 Jahre
Befragung anonym	48,6 % (35/72)	47,9 % (34/71)	63,5 % (33/52)
Befragung nicht-anonym	37,0 % (20/54)	53,9 % (48/89)	47,3 % (26/55)

Für die nicht-anonyme Befragung ist erkennbar, dass sich die relative Dominanz der mittleren Altersgruppe in der Netto-Stichprobe gegenüber der Brutto-Stichprobe (siehe Abbildung 1) u.a. deshalb verstärkt hat, weil die Teilnahmequote der 49- bis 59-Jährigen sehr gering ist. Vergleicht man zudem die altersgruppenspezifischen Response-Raten der anonymen und nicht-anonymen Befragung miteinander, wird deutlich, dass v.a. Angehörige der ältesten und jüngsten Kohorte gegenüber dem nicht-anonymen Befragungsmodus zurückhaltend reagieren, während Personen im Alter zwischen 60 und 70 Jahren dem Anonymitätsmodus indifferent gegenüberstehen bzw. die nicht-anonyme Befragung tendenziell präferieren.<sup>23</sup>

Die Ergebnisse zeigen, dass zumindest im Hinblick auf das Alter der Respondenten nicht die Anonymität, sondern vielmehr die *Nicht*-Anonymität einer Befragung in der Bevölkerung differenziell bewertet wird. Anonyme Befragungen erzeugen im Vergleich zu nicht-anonymen Befragungen entsprechend weniger verzerrte Stichproben und damit qualitativ hochwertigere Ergebnisse. Die Annahme, dass lediglich das Antwortverhalten der Respondenten, nicht aber deren Bereitschaft zur Teilnahme am Survey bzw. zur Beantwortung der gestellten Fragen vom Anonymitätsmodus der Befragung beeinflusst wird (*Bradburn et al.* 1978), erweist sich hier entsprechend als unzutreffend. Dass die Teilnahmeentscheidung einiger Personen

<sup>23</sup> Im Vergleich zur nicht-anonymen Befragung ist die Response-Rate unter anonymen Bedingungen für AG3 um 16,2 % höher, für AG1 um 11,6 % höher und für AG2 um 6,0 % geringer. Ob die festgestellten altersgruppenbezogenen Unterschiede in der Bewertung der Erhebungsanonymität auf Kohorten- und/oder Alterseffekten beruhen, kann aufgrund des Querschnittsdesigns der Befragung nicht entschieden werden. Ebenso können keine empirischen Aussagen zu den Gründen der Altersgruppenunterschiede getroffen werden. Denkbar wäre, dass die erhöhte Vorsicht bezüglich einer Auskunftserteilung bei den ältesten Befragten v.a. aufgrund des Themas „Kriminalität“ und generationen- bzw. alterstypischer sozialer Ängste zustandekommt. Für die jüngste Kohorte kann hier vermutet werden, dass diese besonders medienaffin bzw. tagespolitisch informiert ist und deshalb v.a. aufgrund aktueller gesellschaftlicher Skepsis in Bezug auf die Sicherheit persönlicher Daten besonderen Wert auf Anonymität legt. Beide Erklärungsmuster könnten für Angehörige der mittleren Altersgruppe aufgrund ihres Alters bzw. ihrer Generationenzugehörigkeit weniger zutreffen und entsprechend deren vergleichsweise neutrale Haltung gegenüber dem Anonymitätsmodus der Befragung erklären.

durchaus davon abhängt, ob Anonymität gewährleistet wird, zeigt sich neben einer geringfügig höheren Rücklaufquote unter anonymen Bedingungen auch in der unterschiedlichen Zusammensetzung der resultierenden Netto-Stichproben beider Anonymitätsmodi.

Im Folgenden wird nun das *inhaltliche* Antwortverhalten der befragten Personen in Abhängigkeit von der Erhebungsanonymität genauer betrachtet.

## 5.2 Auswirkungen der Befragungsanonymität auf das Antwortverhalten

Zunächst wurde das Ausmaß der Antwortausfälle über den gesamten Fragebogen in beiden Gruppen miteinander verglichen. Mit einer durchschnittlichen Item-Nonresponse-Rate von 10,3 % weist die anonyme Befragung gegenüber der nicht-anonymen (mit einem entsprechenden Wert von 11,0 %) insgesamt eine geringfügig geringere, aber erneut nicht signifikante<sup>24</sup> Antwort-Verweigerungsquote auf.<sup>25</sup> Die zu beobachtenden Unterschiede beider Befragungsmodi lassen sich im Wesentlichen auf Unterschiede im Ausmaß fehlender Auskünfte *bei der Delinquenzabfrage* zurückführen. Vergleiche der Antwortausfälle beider Befragungsvarianten für insgesamt 42 Abfragen (für 14 Delikte je drei Referenzzeiträume) zeigen, dass der Item-Nonresponse (INR) in der anonymen Befragung abgesehen von drei Items<sup>26</sup> durchgängig geringer ausfällt als in der nicht-anonymen Befragung. Zwar sind diese Unterschiede in den meisten Fällen geringfügig und nur in sechs von 42 Vergleichen, d.h. in 14,3 % der Fälle, statistisch signifikant.<sup>27</sup> In der Tendenz sprechen die Ergebnisse jedoch insgesamt dafür, auch These 2 vorläufig beizubehalten.

In einem nächsten Schritt wurden 91 Fragen bzw. Items, die heikle Informationen im weiter oben benannten Sinne erheben, auf Unterschiedlichkeit der Antwortmuster überprüft. Die betrachteten Fragen bzw. Items beziehen sich überwiegend auf

---

<sup>24</sup>  $T = 0,62$ ;  $p_{2\text{-seitig}} = 0,54$ ;  $df = 194$ .

<sup>25</sup> In der anonymen Befragung fehlen durchschnittlich 20,9 Auskünfte pro Person und Bogen, in der nicht-anonymen fehlen im Mittel 22,1 Antworten.

<sup>26</sup> Bei drei (von 42) Delinquenz-Items ist der INR in der anonymen Befragung gegenüber der nicht-anonymen geringfügig höher. Statistisch signifikant sind diese Unterschiede nicht.

<sup>27</sup> Insgesamt unterscheiden sich die INR-Raten bei je zwei von 14 Delinquenz-Items pro Referenzzeitraum signifikant voneinander. Für folgende Items sind die Unterschiede signifikant ( $p_{2\text{-seitig}} \leq 0,1$ ): Bedrohung/Erpressung: Lebensspanne ( $\Phi = 0,15$ ;  $p = 0,035$ ); Steuerbetrug: Lebensspanne ( $\Phi = 0,13$ ;  $p = 0,069$ ); Sachbeschädigung: seit 50. Geb. ( $\Phi = 0,13$ ;  $p = 0,069$ ); Bedrohung/Erpressung: seit 50. Geb. ( $\Phi = 0,15$ ;  $p = 0,035$ ); Ladendiebstahl: letzte 12 Mon. ( $\Phi = 0,13$ ;  $p = 0,069$ ); Bedrohung/Erpressung: letzte 12 Mon. ( $\Phi = 0,15$ ;  $p = 0,035$ ).

prospektive und retrospektive Angaben zu eigenem delinquenten Verhalten<sup>28</sup> (76 Variablen), aber auch auf Einstellungen gegenüber Gesetzen und Gesetzesübertretungen<sup>29</sup> (15 Variablen). Je nach Messniveau der betrachteten Variablen wurden Unterschiede im Antwortverhalten mittels Kreuztabellierung bzw. T-Test für unabhängige Stichproben analysiert. Insgesamt zeigten sich bei lediglich drei der geprüften Variablen (3,3 % aller Tests) in Abhängigkeit der Befragungsanonymität signifikant unterschiedliche Antwortverteilungen ( $p_{2\text{-seitig}} \leq 0,05$ ). Da angesichts des äußerst geringen Anteils signifikanter Antwortunterschiede ein zufälliges Zustandekommen der Ergebnisse nicht ausgeschlossen werden kann, wird hier auf deren nähere Darstellung und inhaltliche Interpretation verzichtet. Die Test-Ergebnisse legen zunächst nahe, dass das inhaltliche Antwortverhalten bei der Abfrage heikler Themen nicht von der Erhebungsanonymität beeinflusst wird.

Ergänzend wurde nun anhand eines inhaltlich bedeutsamen Beispiels geprüft, inwiefern sich Auswirkungen der Anonymität auf bivariater Ebene, etwa in Antwortkombinationen, zeigen. Es wurde der Einfluss der Befragungsanonymität auf den Zusammenhang zwischen Norminternalisierung und Delinquenzbereitschaft untersucht. Die Erhebung von Norminternalisierung und Delinquenzbereitschaft erfolgte in vorliegendem Fragebogen im Rahmen eines faktoriellen Designs.<sup>30</sup> Es wurden den Respondenten kurze Geschichten (sog. Vignetten oder Szenarios) dargeboten, in denen eine fiktive Person (hier Frau Schmidt genannt) unter je veränderten persönlichen und/oder situativen Umständen eine kriminelle Handlung begeht. Der Fragebogen enthält insgesamt sechs Vignetten, wovon sich die ersten drei auf Ladendiebstahl und die letzten drei auf Versicherungsbetrug (speziell Übertreibung der Schadenssumme) beziehen.<sup>31</sup> Den Vignetten schließen sich fünf Fragen (sog.

---

<sup>28</sup> Analysiert wurden Angaben zur Delinquenzbereitschaft, zur eigenen Delinquenz (14 Delikte, je Prävalenz und Inzidenz) in Bezug auf drei Referenzzeiträume („Lebensspanne“, „seit dem 50. Geburtstag“, „in den letzten 12 Monaten“) sowie zur polizeilichen Registrierung in verschiedenen Lebensabschnitten.

<sup>29</sup> Analysiert wurden Angaben zur moralischen Bewertung verschiedener strafbarer Verhaltensweisen sowie zu allgemeinen Einstellungen gegenüber Gesetzen.

<sup>30</sup> Zum Faktoriellen Surveyansatz (engl. Factorial Survey Approach: FSA) siehe einführend *Rossi & Nock* 1982. Der FSA weist in vielerlei Hinsicht Ähnlichkeiten mit der v.a. im deutschsprachigen Raum bekannten Facettentheorie (*Borg* 1999) auf.

<sup>31</sup> Die Vignetten bilden sechs inhaltliche Dimensionen bzw. Faktoren ab, die jeweils zwei mögliche Ausprägungen bzw. Levels haben. Das sog. Faktoruniversum, d.h. die Anzahl aller möglichen Faktor-/Level-Kombinationen beträgt damit pro Delikttyp 64 Vignetten (Kartesisches Produkt:  $2^6$ ). Für beide Delikttypen (Ladendiebstahl, Versicherungsbetrug) ergeben sich insgesamt 128 verschiedene Vignetten. Um das Pretest-Design nicht komplizierter als nötig zu gestalten, wurden im Pretest-Fragebogen lediglich je drei Vignetten für beide Delikttypen aus dem je entsprechenden Faktoruniversum, insgesamt also sechs Vignetten, zufällig ausgewählt. Diese waren für alle Befragten fix. In der Haupterhebung wurden sämtliche Faktor-/Level-Kombinationen der Vignetten – allerdings verteilt auf 32 verschiedene Fragebogenversionen, die an je 111 bzw. 112 Personen versendet wurden – zur Bewertung vorgelegt.

Rating-Tasks) an, welche die Bewertung der jeweiligen Situation und Handlung sowie die Einschätzung der eigenen Delinquenzbereitschaft beinhalten. Die Abfrage von Norminternalisierung<sup>32</sup> bzw. Delinquenzbereitschaft<sup>33</sup> erfolgte jeweils am Anfang bzw. am Ende des vignettenbezogenen Frageblocks.<sup>34</sup> In einem ersten Schritt wurden für alle sechs Vignetten univariate OLS-Regressionen mit Delinquenzbereitschaft als abhängiger und Norminternalisierung als unabhängiger Variable berechnet. Anschließend wurde ermittelt, ob die jeweils unstandardisierten Regressionskoeffizienten beider Gruppen signifikant voneinander abweichen.

Aus Tabelle 4 geht zunächst einmal hervor, dass der Zusammenhang beider Variablen für nahezu alle Vignetten in beiden Befragungsvarianten zwar die gleiche und erwartete negative Richtung aufweist, d.h. je stärker die Norminternalisierung ausgeprägt ist, desto geringer ist die Delinquenzbereitschaft. Die Zusammenhangsstärke variiert jedoch deutlich zwischen den experimentellen Konditionen (sowie zwischen den Vignetten innerhalb einer Kondition, was hier jedoch weniger von Interesse ist).

Die Ergebnisse zeigen weiter, dass sich die Zusammenhangsstärken in Abhängigkeit der Befragungsanonymität bei vier der sechs untersuchten Vignetten (Vignetten 1 und 3 bis 5), d.h. in zwei Drittel der Fälle, auf dem 5%-Niveau signifikant voneinander unterscheiden.<sup>35</sup> Dabei treten die deutlich stärkeren Zusammenhänge zwischen Norminternalisierung (NI) und Delinquenzbereitschaft (DB) durchgehend in der nicht-anonymen Befragung auf.<sup>36</sup> Diese Resultate sind v.a. deshalb interes-

---

32 „Wie bewerten Sie das Verhalten von Frau Schmidt **unter den beschriebenen Umständen**? Ist das für Sie ein schlimmes oder weniger schlimmes Verhalten?“ (Fettdruck im Original-Fragebogen, erhoben für 6 Vignetten); Antwortvorgaben: „überhaupt nicht schlimm“ (1) ... „ganz besonders schlimm“ (10).

33 „Stellen Sie sich bitte einmal vor, Sie selbst wären in der Lage von Frau Schmidt. Wie wahrscheinlich würden Sie dann auch so handeln wie Frau Schmidt in der Geschichte x?“ (erhoben für 6 Vignetten); Antwortvorgaben: „völlig unwahrscheinlich“ (1) ... „völlig wahrscheinlich“ (10).

34 Die dazwischen liegenden Fragen, die Rational-Choice- und Lernvariablen erheben, fungieren gleichzeitig als „Puffer“, um Einflüsse kognitiver Konsistenzbestrebungen auf die Antworten zu Norminternalisierung und Delinquenzbereitschaft abzuschwächen. Um Effekte des Fragekontexts (bzw. der Frageposition) auf das Antwortverhalten der Respondenten zu untersuchen, wurde die Reihenfolge der Abfrage von Norminternalisierung und krimineller Verhaltensintention im Methodenexperiment variiert. Bei je der Hälfte der Befragten (d.h. bei je 200 Personen) wurde zuerst die Norminternalisierung und zuletzt die eigene Delinquenzbereitschaft bzw. zuerst die Delinquenzbereitschaft und zuletzt die Norminternalisierung erfragt; Näheres zu dieser Teilstudie: *Kunz & Grundies 2011*.

35 Bei Vignette 6 wird das 5%-Signifikanzniveau nur knapp verfehlt; die Unterschiede sind hier mit  $p = 0,06$  statistisch signifikant.

36 Inhaltlich bedeutet dieses Ergebnis, dass die Auskünfte der anonym Befragten zur moralischen Beurteilung des Verhaltens von Frau Schmidt (NI) und zur Wahrscheinlich-

sant, weil sich in der univariaten Analyse der Antworten zu Norminternalisierung und Delinquenzbereitschaft (siehe weiter oben) keine Unterschiede in Abhängigkeit der Befragungsanonymität gezeigt hatten. Folgt man den Annahmen, dass a) ein stärkerer negativer Zusammenhang beider Variablen (auch) höhere Antwortkonsistenz ausdrückt und dass sich b) darin zugleich (auch) ein höheres Maß sozial erwünschten Antwortverhaltens widerspiegelt (Scheuch 1974, S. 90; Hafermalz 1976, S. 229), dann legen die Ergebnisse der bivariaten Analyse (im Gegensatz zu denen der univariaten Analyse) den Schluss nahe, dass die Teilnehmer unter anonymen Bedingungen in geringerem Ausmaß sozial erwünscht antworten.

*Tabelle 4: Zusammenhang zwischen Delinquenzbereitschaft und Norminternalisierung in Abhängigkeit von Vignette und Befragungsanonymität\* (OLS-Regression, unstandardisierte Regressionskoeffizienten B, in Klammern: SE<sub>B</sub>)*

	nicht-anonym	anonym
Vignette 1	-0,45 (0,09)	0,06 (0,14)
Vignette 2	-0,29 (0,11)	-0,23 (0,12)
Vignette 3	-0,31 (0,09)	-0,02 (0,12)
Vignette 4	-0,47 (0,10)	-0,18 (0,09)
Vignette 5	-0,48 (0,09)	-0,23 (0,09)
Vignette 6	-0,45 (0,08)	-0,24 (0,08)

\* Grau unterlegte Zellen markieren signifikant unterschiedliche Regressionskoeffizienten. Die Signifikanz der Differenz wurde mittels folgender Teststatistik bestimmt:

$$z = \frac{B_1 - B_2}{\sqrt{SE_{B_1}^2 + SE_{B_2}^2}}; \text{ wobei } B_i = \text{unstand. Regressionskoeffizient und } SE_{B_i} = \text{Standardfehler des unstand. Regressionskoeffizienten.}$$

Für z-Werte  $\geq 11,961$  (entspricht  $p \leq 0,05$ ) wurde die Nullhypothese  $H_0: B_1 = B_2$  abgelehnt.

Wie weiter vorn gezeigt wurde, sind die Teilnehmer der anonymen Befragung im Vergleich zu denen der nicht-anonymen Befragung durchschnittlich jünger. Dies ist u.a. darauf zurückzuführen, dass der Effekt der Befragungsanonymität auf die Teilnahmebereitschaft vom Alter bzw. von der Altersgruppenzugehörigkeit der kontaktierten Personen moderiert wird (Abschnitt 5.1). Es wäre deshalb nun denkbar, dass die beobachteten Unterschiede des Zusammenhangs zwischen NI und DB (teilwei-

keit, sich unter gleichen Umständen selbst so zu verhalten (DB), entgegen dissonanztheoretischer Erwartungen im Vergleich zu den entsprechenden Auskünften der nicht-anonym Befragten häufiger so ausfielen, dass sie das beschriebene kriminelle Verhalten zwar (eher) schlimm finden, sich selbst aber ähnlich verhalten würden und umgekehrt. Dieses scheinbar widersprüchliche Antwortverhalten, das bei einem Teil der Befragten zu beobachten ist, kann z.B. unter Rückgriff auf normtheoretische (Geiger 1964; Popitz 1980; 2003 [1968]) bzw. neutralisationstheoretische Überlegungen (Sykes & Matza 1957) erklärt werden.

se oder gänzlich) vom Alter der Befragten abhängen und ergo (teilweise oder gänzlich) auf entsprechenden Unterschieden in der Stichprobenzusammensetzung beruhen. In einem letzten Schritt wurde dies überprüft, indem das Alter der Respondenten zusätzlich zur Norminternalisierung in sämtliche vignettenspezifischen OLS-Regressionsgleichungen einbezogen wurde und anschließend die Unterschiede beider Samples im Hinblick auf die jeweils resultierenden unstandardisierten Regressionskoeffizienten der Norminternalisierung erneut mithilfe der angegebenen Teststatistik bewertet wurden. Die Analysen zeigen zunächst für beide Samples und – abgesehen von einer Ausnahme<sup>37</sup> – für alle Vignetten, dass das Alter der Befragten keinen unabhängigen statistisch bedeutsamen Einfluss auf die Delinquenzbereitschaft ausübt. Weiter ist festzustellen, dass der Unterschied beider Samples in Bezug auf den Zusammenhang zwischen Delinquenzbereitschaft und Norminternalisierung für Vignette 3 (Ladendiebstahl), wenn das Alter der Befragten berücksichtigt wird, nun nicht mehr mit  $p = 0,05$  sondern lediglich mit  $p = 0,054$  (entspricht  $z = -1,93$ ) signifikant ist. Alle übrigen Resultate (siehe Tabelle 4) werden auch unter Kontrolle des Alters nahezu unverändert reproduziert. Alles in allem ist der Einfluss des Alters damit so marginal, dass weiter davon auszugehen ist, dass die berichteten Ergebnisse zur Auswirkung der Befragungsanonymität auf den Zusammenhang zwischen Norminternalisierung und Delinquenzbereitschaft keine Artefakte, sondern inhaltlich valide und robuste Befunde darstellen.

Abschließend kann Folgendes festgehalten werden: Zwar unterscheiden sich die Respondenten beider Befragungsvarianten wie erwartet hinsichtlich ihrer Teilnahmebereitschaft (These 1) sowie hinsichtlich des Ausmaßes an Item-Nonresponse in Bezug auf heikle Fragen voneinander (These 2). Was das inhaltliche Antwortverhalten betrifft (These 3) sind die Ergebnisse allerdings uneindeutig. Auf univariater Ebene zeigt sich entgegen der theoretischen Annahmen kein nennenswerter Effekt der Erhebungsanonymität. Anhand bivariater Analysen konnten die erwarteten Unterschiede in Form stärkerer Antwortkonsistenz bei nicht-anonymer Befragung allerdings klar festgestellt werden. Insgesamt sind also Auswirkungen der Befragungsanonymität auf das inhaltliche Antwortverhalten der Respondenten zu beobachten, wenngleich in geringerem Ausmaß und in anderer Form (nämlich eher subtil) als erwartet.

Als mögliche Erklärungen für das relativ geringe Ausmaß der Antwortunterschiede kommen im Wesentlichen jene Annahmen in Betracht, die bereits zur Begründung der nur schwach unterschiedlichen Teilnahmequote diskutiert wurden. So erscheint es u.a. denkbar, dass bereits der hohe Anonymitätsgrad postalischer Befragungen

---

<sup>37</sup> Lediglich in einer von insgesamt 12 Analysen übt das Alter der Respondenten einen signifikanten (negativen) Einfluss auf die berichtete Delinquenzbereitschaft aus (d.h. je älter die Befragten, desto geringer ist deren Delinquenzbereitschaft). Dies gilt nur für Vignette 6 unter *nicht-anonymen* Befragungsbedingungen: Beta = -0,20;  $p = 0,02$ ;  $n = 93$ .

und/oder die Vergabe monetärer Incentives potenzielle Einflüsse der Befragungsanonymität auf das Antwortverhalten hemmen bzw. überlagern (vgl. die Argumentation in Abschnitt 5.1).

## **6. Zusammenfassung und methodologische Empfehlungen**

Die vorliegende Studie untersuchte mithilfe eines experimentellen Designs, inwiefern sich die Erhebungsanonymität postalischer Befragungen zu selbstberichteter Kriminalität auf die Teilnahmebereitschaft und das Antwortverhalten der Respondenten auswirkt. Unter anonymen Befragungsbedingungen wurde eine leicht erhöhte Rücklaufquote (Differenz ca. 5 %) beobachtet. Es konnte ferner gezeigt werden, dass die Bereitschaft, an der nicht-anonymen Befragung teilzunehmen, von der Zugehörigkeit zu bestimmten Altersgruppen abhängt, während die Teilnahmebereitschaft unter anonymen Bedingungen vom Alter der Befragten vergleichsweise unbeeinflusst bleibt. Anonyme Befragungen erzeugen im Vergleich zu nicht-anonymen Befragungen damit weniger verzerrte Stichproben und sind deshalb besser für eine Verallgemeinerung stichprobenbasierter Befunde geeignet.

Die Untersuchung des Einflusses der Befragungsanonymität auf das Antwortverhalten der Respondenten erbrachte gemischte Befunde: Auf univariater Ebene wurden keine Hinweise darauf gefunden, dass sich das inhaltliche Antwortverhalten der Respondenten in Abhängigkeit der Befragungsanonymität unterscheidet. Allerdings ist das Gesamtausmaß des Item-Nonresponse in der anonymen Befragung minimal geringer als in der nicht-anonymen. Diese Unterschiede beruhen im Wesentlichen darauf, dass Auskünfte zur eigenen Delinquenz unter nicht-anonymen Erhebungsbedingungen tendenziell häufiger unbeantwortet bleiben.

Die bivariaten Analysen belegen, dass Zusammenhänge zwischen Variablen in theoretisch erwarteter Weise von der Erhebungsanonymität beeinflusst werden können. Exemplarisch wurde gezeigt, dass die Kombination der Antworten auf Fragen zur Norminternalisierung und Delinquenzbereitschaft in zwei Drittel der Testsituationen von der Befragungsanonymität abhängen. Im Vergleich zur nicht-anonymen Befragung wurden in der anonymen Befragung durchweg schwächere Variablenzusammenhänge beobachtet, was hier als weniger sozial erwünschtes Antwortverhalten interpretiert wird.

Insgesamt zeigt die vorliegende Untersuchung, dass anonyme Befragungen im Vergleich zu nicht-anonymen durch eine höhere Stichprobenqualität, geringere Unit- und Item-Nonresponse-Raten sowie durch weniger sozial erwünschtes Antwortverhalten charakterisiert sind. Allerdings fallen die beobachteten Unterschiede anonymer und nicht-anonymer Befragungen je für sich genommen eher schwach aus.

Aus rein methodologischer Sicht empfehlen die Ergebnisse der durchgeführten Untersuchung bei postalischen Befragungen zu heiklen Themen, speziell zu selbst-

berichteter Kriminalität, die vollständige Anonymität der Befragten zu gewährleisten und entsprechend auf eine Paginierung der Fragebögen zu verzichten. Dies gilt insbesondere, wenn anhand der Daten genaue Schätzungen des Ausmaßes bestimmter Einstellungen oder Verhaltensweisen in der Grundgesamtheit vorgenommen werden sollen. Da die anonyme Befragungsvariante aber mit einem finanziellen Mehraufwand und forschungspraktischen Einschränkungen verbunden ist,<sup>38</sup> sollten die Vor- und Nachteile anonymer und nicht-anonymer Befragung für jede Untersuchung vor dem Hintergrund ihrer jeweiligen Ziele und Ausrichtung (Zielgruppe, Themen etc.) kritisch gegeneinander abgewogen werden. Sofern bei der Erhebung sensibler Informationen Identifikationsnummern verwendet werden, sollten diese jedoch im Anschreiben erwähnt und ihre Funktion erläutert werden. „Das Gefühl der Anonymität auf Seiten der Befragten ist unter diesen Umständen durch andere geeignete Maßnahmen sicherzustellen“ (*Reuband* 2011, S. 31). Hierzu zählen z.B. die Zusicherung einer vertraulichen und anonymen Datenauswertung sowie die Durchführung der Studie durch vertrauenswürdige Personen bzw. Institutionen (Sponsorship-Effekt).

Da die Datenbasis der Studie sehr eingeschränkt ist und einige Effekte daher nur tendenziell beobachtbar sind bzw. keine verlässlichen Aussagen zu deren statistischer (In-)Signifikanz getroffen werden können, sollten die hier referierten Befunde zunächst durch weitere Experimente (mit deutlich größerer Stichprobe) abgesichert werden. Es erscheint lohnenswert, im Rahmen solcher Studien auch zu untersuchen, ob und inwiefern Effekte der Erhebungsanonymität auf das Teilnahme- und Antwortverhalten der Respondenten durch andere methodologische Aspekte (z.B. Vergabe von Incentives, Vertrauenswürdigkeit der Forschungsinstitution) moderiert werden.

---

<sup>38</sup> Hierzu zählen etwa: erhöhte Kosten durch den zusätzlichen Postkartenversand, keine Verknüpfung der Respondentenauskünfte mit früheren bzw. späteren Auskünften oder mit anderen Daten (z.B. für eine Validierung) möglich.

## Literatur

- Albaum, G.* (1987): Do source and anonymity affect mail survey results? *Journal of the Academy of Marketing Science* 15, S. 74–81.
- Arzheimer, K. & Klein, M.* (1998): Die Wirkung materieller Incentives auf den Rücklauf einer schriftlichen Panelbefragung. *ZA-Information* 43, S. 6–43.
- Ash, P. & Abramson, E.* (1952): The effect of anonymity on attitude questionnaire response. *Journal of Abnormal Psychology* 47, S. 722–723.
- Baurmann, M., Hermann, D., Störzer, U. & Streng, F.* (1991): Telefonische Befragung von Kriminalitätsoffern: ein neuer Weg ins Dunkelfeld? *Monatsschrift für Kriminologie und Strafrechtsreform*, S. 159–173.
- Becker, G. & Bakal, D.A.* (1970): Subject anonymity and motivational distortion in self-report data. *Journal of Clinical Psychology*, S. 207–209.
- Becker, R. & Günther, R.* (2004): Selektives Antwortverhalten bei Fragen zum delinquenten Handeln. Eine empirische Studie über die Wirksamkeit der „sealed envelope technique“ bei selbst berichteter Delinquenz mit Daten des ALLBUS 2000. *ZUMA-Nachrichten* 52, S. 39–59.
- Becker, R., Imhof, R. & Mehlkop, G.* (2007): Die Wirkung monetärer Anreize auf den Rücklauf bei einer postalischen Befragung und die Antworten auf Fragen zur Delinquenz. *Methoden Daten Analysen* 1/2, S. 131–160.
- Becker, R. & Mehlkop, G.* (2011): Effects of prepaid monetary incentives on mail survey response rates and on self-reporting about delinquency – empirical findings. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique* 111, S. 5–25.
- Benson, L.E.* (1941): Studies in secret-ballot technique. *Public Opinion Quarterly* 5, S. 79–82.
- Berman, J., McCombs, H. & Boruch, R.F.* (1977): Notes on the contamination Method: Two small experiments in assuring confidentiality of responses. *Sociological Methods and Research* 6, S. 45–62.
- Blau, P.M.* (1964): *Exchange and Power in Social Life*. New York.
- Blumberg, H.H., Fuller, C. & Hare, P.A.* (2004): Response rates in postal surveys. *Public Opinion Quarterly* 38, S. 113–123.
- Borg, I.* (1999): *Grundlagen und Ergebnisse der Facettentheorie*. Bern.
- Bradburn, N.M., Sudman, S., Blair, E. & Stocking, C.* (1978): Question threat and response bias. *Public Opinion Quarterly* 42, S. 221–234.
- Brennan, M.* (1992): The effect of a monetary incentive on mail survey response rates: New data. *Journal of the Market Research Society* 34/2, S. 173–177.
- Campbell, M.J. & Waters, W.E.* (1990): Does anonymity increase response rate in postal questionnaire surveys about sensitive subjects? A randomised trial. *Journal of Epidemiology and Community Health* 44, S. 75–76.

- Cohen, J., Cohen, P., West, S.G. & Aiken, L.S.* (2003): Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences. Mahwah, NJ.
- Cohen, J.* (1977): Statistical Power Analysis for the Behavioral Sciences. New York.
- Deutsche Forschungsgemeinschaft (1999): Qualitätskriterien der Umfrageforschung. Berlin.
- Diekmann, A. & Jann, B.* (2001): Anreizformen und Ausschöpfungsquoten bei postalischen Befragungen. Eine Prüfung der Reziprozitätshypothese. ZUMA-Nachrichten 48, S. 18–27.
- Dillman, D.A.* (2000): Mail and Internet Surveys: The Tailored Design Method. New York.
- Dillman, D.A.* (1978): Mail and Telephone Surveys. New York.
- Dillman, D.A., Smyth, J.D. & Christian, L.M.* (2009): Internet, Mail, and Mixed-Mode Surveys. The Tailored Design Method. Hoboken, NJ.
- Epperson, W.V. & Peck, R.C.* (1977): Questionnaire response bias as a function of respondent anonymity. Accident Analyses and Prevention 9, S. 249–256.
- Esser, H.* (1986): Können Befragte lügen? Zum Konzept des „wahren Wertes“ im Rahmen der handlungstheoretischen Erklärung von Situationseinflüssen bei der Befragung. Kölner Zeitschrift für Soziologie und Sozialpsychologie 38, S. 314–336.
- Fischer, R.P.* (1946): Signed versus unsigned personal questionnaires. Journal of Applied Psychology 30, S. 220–225.
- Fuller, C.H.* (1974): Effect of anonymity on return rate and response bias in a mail survey. Journal of Applied Psychology 59, S. 292–296.
- Futrell, Ch.M.* (1981): Effects of signed versus unsigned attitude questionnaires. Journal of the Academy of Marketing Science 9, S. 93–98.
- Futrell, Ch.M. & Swan, J.E.* (1977): Anonymity and response by salespeople to a mail questionnaire. Journal of Marketing Research 14, S. 611–616.
- Geiger, Th.* (1964): Vorstudien zu einer Soziologie des Rechts. Neuwied a. Rhein, Berlin.
- Goode, W.H. & Halt, P.K.* (1965): Die schriftliche Befragung, in: König, R. (Hrsg.), Formen – Technik – Auswertung (Praktische Sozialforschung I). Köln, Berlin, S. 161–177.
- Gouldner, A.W.* (1960): The norm of reciprocity. American Sociological Review 25, S. 161–178.
- Hafermalz, O.* (1976): Schriftliche Befragung – Möglichkeiten und Grenzen. Wiesbaden.
- Hermann, D.* (2003): Werte und Kriminalität. Konzeption einer allgemeinen Kriminalitätstheorie. Wiesbaden.
- Hindelang, M.J., Hirschi, T. & Weis, J.G.* (1979): Correlates of delinquency. The illusion of discrepancy between self-report and official measures. American Sociological Review 44, S. 995–1014.
- Hippler, H.-J., Schwarz, N. & Singer, E.* (1990): Der Einfluss von Datenschutzzusagen auf die Teilnahmebereitschaft an Umfragen. ZUMA-Nachrichten 27, S. 54–67.
- Homans, G.C.* (1961): Social Behavior. Its Elementary Forms. New York.

- James, J.M. & Bolstein, R.* (1990): The effect of monetary incentives and follow-up mailings on the response rate and response quality in mail surveys. *Public Opinion Quarterly* 54, S. 346–361.
- Jones, W.* (1979): Generalizing mail survey inducement methods: Population interactions with anonymity and sponsorship. *Public Opinion Quarterly* 43, S. 102–111.
- Kanuk, L. & Berenson, C.* (1975): Mail surveys and response rates: A literature review. *Journal of Marketing Research* 12, S. 440–453.
- Kerschke-Risch, P.* (1993): Gelegenheit macht Diebe – Doch Frauen klauen auch. Opladen.
- King, F.W.* (1970): Anonymous versus identifiable questionnaires in drug usage surveys. *American Psychologist* 25/10, S. 982–985.
- Kunz, F.* (2014 [im Erscheinen]): Moral und (Un-)Rechtsempfinden im gesellschaftlichen Wandel. Ein empirischer Vergleich von Generationen 50+. Berlin.
- Kunz, F.* (2013a): Kriminalität älterer Menschen. Ergebnisse einer postalischen Befragung. *Geriatric Report* 8/1, S. 8–14.
- Kunz, F.* (2013b [im Erscheinen]): Kriminalität älterer Menschen. Beschreibung und Erklärung auf der Basis von Selbstberichtsdaten. Berlin.
- Kunz, F.* (2011): Kriminelles Verhalten und polizeiliche Registrierung: Selbstberichte von Menschen im höheren Lebensalter. *Zeitschrift für Gerontologie und Geriatrie* 44/1, S. 55–65.
- Kunz, F. & Grundies, V.* (2011): Zur Messung von Norminternalisierung und Delinquenzbereitschaft in Befragungen: Beeinflusst der Fragekontext das Antwortverhalten der Respondenten? *Monatsschrift für Kriminologie und Strafrechtsreform* 94/5, S. 345–363.
- Kury, H.* (1994): Zum Einfluss der Art der Datenerhebung auf die Ergebnisse einer Umfrage. *Monatsschrift für Kriminologie und Strafrechtsreform*, S. 22–33.
- Mauss, M.* (1985 [1925]): Die Gabe. Form und Funktion des Austauschs in archaischen Gesellschaften. Frankfurt.
- McDaniel, S.W. & Rao, C.P.* (1980): The effect of monetary inducement on mailed questionnaire response quality. *Journal of Marketing Research* 17, S. 265–268.
- Mehlhop, G. & Becker, R.* (2007): Zur Wirkung monetärer Anreize auf die Rücklaufquote in postalischen Befragungen zu kriminellen Handlungen. *Methoden Daten Analysen* 1/1, S. 5–24.
- Oberwittler, D., Blank, T., Köllisch, T. & Naplava, T.* (2001): Soziale Lebenslagen und Delinquenz von Jugendlichen. Ergebnisse der MPI-Schulbefragung 1999 in Freiburg und Köln. Freiburg im Breisgau.
- Olson, W.C.* (1936): The waiver of signature in personal reports. *Journal of Applied Psychology* 20, S. 442–450.
- Ong, A. & Weiss, D.J.* (2000): The impact of anonymity on responses to sensitive questions. *Journal of Applied Social Psychology* 30, S. 1691–1708.
- Popitz, H.* (1980): Die normative Konstruktion von Gesellschaft. Tübingen.

- Popitz, H.* (2003 [1968]): Über die Präventivwirkung des Nichtwissens. Berlin.
- Porst, R.* (1999): Thematik oder Incentives? Zur Erhöhung der Rücklaufquoten bei postalischen Befragungen. ZUMA-Nachrichten 45, S. 72–87.
- Prüfer, P. & Rexroth, M.* (2000): Zwei-Phasen-Pretesting. ZUMA-Arbeitsbericht 2000/08. ZUMA. Mannheim.
- Reips, U.-D. & Franek, L.* (2004): Mitarbeiterbefragungen per Internet oder Papier? Der Einfluss von Anonymität, Freiwilligkeit und Alter auf das Antwortverhalten. Wirtschaftspsychologie 1, S. 67–83.
- Reuband, K.-H.* (2011): Freiwillige De-Anonymisierung in anonymen postalischen Befragungen. Wie sehr unterscheiden sich die Befragten mit und ohne De-Anonymisierungsbereitschaft in ihren sozialen Merkmalen, ihrer Devianz und Devianzbereitschaft? Monatsschrift für Kriminologie und Strafrechtsreform 94/1, S. 17–33.
- Reuband, K.-H.* (2006): Postalische Befragung alter Menschen. Kooperationsverhalten, Beantwortungsstrategien und Qualität der Antworten. ZA-Informationen 59, S. 100–127.
- Reuband, K.-H.* (2002): Mail and telephone surveys. Their impact on responses, in: Blasius, J., Hox, J., de Leeuw, J.E. & Schmidt, P. (eds.), Social Science Methodology in the New Millenium [CD-Rom]. Opladen.
- Reuband, K.-H.* (2000): Telefonische und postalische Umfragen in Ostdeutschland. Auswirkungen auf die soziale Zusammensetzung und das Antwortverhalten, in: Hüfken, V. (Hrsg.), Methoden in Telefonumfragen. Opladen, S. 191–215.
- Reuband, K.-H.* (1999): Anonyme und nicht-anonyme postalische Bevölkerungsbefragungen. Planung und Analyse 1, S. 56–59.
- Rossi, P.H. & Nock, S.L.* (eds.) (1982). Measuring Social Judgements. The Factorial Survey Approach. Beverly Hills, CA.
- Scheuch, E.K.* (1974): Das Interview in der empirischen Sozialforschung, in: König, R. (Hrsg.), Handbuch der empirischen Sozialforschung, 2. Stuttgart, S. 66–190.
- Schneekloth, U. & Leven, I.* (2003): Woran bemisst sich eine „gute“ allgemeine Bevölkerungsumfrage? Analysen zu Ausmaß, Bedeutung und zu den Hintergründen von Non-response in zufallsbasierten Stichprobenerhebungen am Beispiel des ALLBUS. ZUMA-Nachrichten 53, S. 16–57.
- Schnell, R., Hill, P.B. & Esser, E.* (1995): Methoden der empirischen Sozialforschung. München, Wien.
- Singer, E., von Thurn, D.R. & Miller, E.R.* (1995): Confidentiality assurances and responses. Public Opinion Quarterly 59, S. 66–77.
- Singer, E.* (1998): Incentives for survey participation: Research on intended and unintended consequences. ZUMA-Nachrichten 42, S. 7–28.
- Singer, E., van Hoewyk, J. & Maher, M.P.* (1998): Does the payment of incentives create expectation effects? Public Opinion Quarterly 62, S. 152–164.
- Stadtmüller, S. & Porst, R.* (2005): Zum Einsatz von Incentives bei postalischen Befragungen. ZUMA How-to-Reihe Nr. 14.

- Stadtmüller, S.* (2009): Rücklauf gut, alles gut? Zu erwünschten und unerwünschten Effekten monetärer Anreize bei postalischen Befragungen. *Methoden Daten Analysen* 3/2, S. 167–186.
- Stoop, I.A.L.* (2005): *The Hunt for the Last Respondent. Nonresponse in Sample Surveys.* The Hague.
- Sykes, G.M. & Matza, D.* (1957): *Techniques of Neutralization: A Theory of Delinquency.* *American Sociological Review* 22/6, S. 664–670.
- Tourangeau, R., Rips, L.J. & Rasinski, K.A.* (2000): *The Psychology of Survey Response.* Cambridge, MA.
- Tyagi, P.K.* (1989): The effects of appeals, anonymity, and feedback on mail survey response patterns from salespeople. *Journal of the Academy of Marketing Science* 17, S. 235–241.
- van de Looij-Jansen, P.M., Goldschmeding, J.E.J. & de Wilde, E.J.* (2006): Comparison of anonymous versus confidential survey procedures: Effects on health indicators in Dutch adolescents. *Journal of Youth and Adolescence* 35, S. 659–665.
- Wildman, R.C.* (1977): Effects of anonymity and social setting on survey responses. *Public Opinion Quarterly* 41, S. 74–79.
- Wolfgang, M.E., Thornberry, T.B. & Figlio, R.M.* (1987): *From Boy to Man, From Delinquency to Crime.* Chicago.
- Wotruba, Th.R.* (1966): Monetary inducements and mail questionnaire research. *Journal of Marketing Research* 3, S. 393–400.
- Yammarino, F.J., Skinner, S.J. & Childers, T.L.* (1991): Understanding mail survey response behavior. *Public Opinion Quarterly* 55, S. 613–639.
- Yu, J. & Cooper, H.* (1983): A quantitative review of research design effects on response rates to questionnaires. *Journal of Marketing Research* 20, S. 36–44.



Forschung zum Einfluss der Erhebungsanonymität auf die Rücklaufquote und das Antwortverhalten in postalischen Befragungen erbrachte bislang widersprüchliche Ergebnisse; für kriminalitätsbezogene Umfragen wurden solche Effekte zudem kaum untersucht. Der vorliegende Beitrag berichtet entsprechende Ergebnisse einer experimentellen Pilotstudie, die im Jahr 2009 im Rahmen einer standardisierten postalischen Befragung unter 49- bis 81-Jährigen zu selbstberichteter Kriminalität im Raum Sübaden/Deutschland durchgeführt wurde.

Die Analysen zeigen, dass die anonyme Befragung im Vergleich zur nicht-anonymen tendenziell durch eine höhere Stichprobenqualität, geringere Unit- und Item-Nonresponse-Raten sowie weniger sozial erwünschtes Antwortverhalten gekennzeichnet ist.

Postalische Befragungen zu heiklen Themen, speziell zu selbstberichteter Kriminalität, sollten daher nach Möglichkeit anonym durchgeführt werden.

# research in brief