

Mahnaktionen in postalischen Befragungen

*Empirische Befunde zu
Auswirkungen auf den
Rücklauf, das Antwort-
verhalten und die Stich-
probenzusammensetzung*

Follow-Ups in Mail Surveys

*Empirical Findings on the
Effects on the Response
Rate, the Responses and
the Sample Composition*

Franziska Kunz

Zusammenfassung

Empirische Befunde zur rücklaufsteigernden Wirkung von Remindern sind konsistent und zahlreich. Nur wenige Studien beschäftigen sich mit dem Einfluss von Nachfassaktionen auf das Antwortverhalten der Respondenten und die soziale Zusammensetzung der Befragtenstichprobe. Vorliegender Beitrag untersucht diese Aspekte anhand der Daten einer im Jahr 2009 im Raum Südbaden durchgeführten postalischen Befragung (N=3.555) unter 49- bis 81-Jährigen zum Thema „Kriminalität im höheren Lebensalter“. Die Analysen ergeben, dass das Ausmaß fehlender Antworten (Item-Nonresponse) mit den Mahnaktionen zunimmt, wobei der Zusammenhang zwischen der Zahl der Nachfassaktionen und der Zahl fehlender Antworten kurvilinear ist. Mahnaktionen rekrutieren v. a. Gruppierungen in die Befragungsstichprobe, die nach einmaligem Anschreiben üblicherweise unterrepräsentiert sind – nämlich Personen, die jünger sind, im ländlichen Raum wohnen sowie einen schwächeren Sozialstatus und geringeres politisches Interesse aufweisen. Mahnaktio-

Abstract

The fact that reminders increase the response rates of mail surveys is empirically documented. However, only a few studies have dealt with the effect of follow-up mailings on both response behaviour and sample composition. This article investigates these aspects using data from a regionally representative mail survey (N=3.555) on crime at advanced ages among 49 to 81 years-olds that was conducted in Südbaden/Baden-Württemberg in 2009. The analyses showed that the use of follow-ups increased the number of missing values (item-nonresponse), with a curvilinear relation between the number of follow-ups and the extent of item-nonresponse. Follow-ups primarily involve people that are usually underrepresented in the first wave, i. e., people who are younger, living in rural areas, having a lower social status and weaker political interest. Using reminders therefore contributes to an enhanced social composition of the net sample. Nevertheless, the respondents of the first mailing do not differ fundamentally from those of the

nen tragen damit zu einer Verbesserung der sozialen Zusammensetzung der Stichprobe bei. Hinsichtlich sozio-demographischer Merkmale sowie psychosozialer Befindlichkeiten unterscheiden sich die Befragten des Erstversands allerdings nicht gravierend von denen der gesamten Befragungsstichprobe. Eine Ausnahme stellt das kriminelle Handeln der Befragten über die Lebensspanne dar: mit zunehmenden Kontaktierungen werden deutlich weniger delinquente Personen in die Befragung einbezogen. Die Antworten der Respondenten zum Befragungsthema stehen demnach in deutlichem Zusammenhang zum Erhebungszeitpunkt. Nachfassaktionen können somit auch hier zur Abschwächung anfänglicher Verzerrungen beitragen. Aus methodischer Sicht sollten Mahnaktionen deshalb prinzipiell in das Erhebungsdesign postalischer Befragungen eingeplant werden. Insgesamt drei Erhebungszeitpunkte erscheinen dabei optimal.

complete survey sample in terms of their social variables as well as their psycho-social characteristics. However, one exception is the respondents' criminal behaviour over the life-course: with increasing numbers of mailings the respondents' delinquency strongly decreases. Obviously, responses on the survey topic are thus correlated with the timing of data collection. This again shows how reminders can help diminishing biases. From a methodological point of view it is therefore recommended to always include follow-up mailings in postal surveys. Three mailings seem to be ideal.

1 Einleitung¹

Im Vergleich zu anderen Erhebungsmodi herrscht gegenüber der postalischen Befragung als wissenschaftlicher Methode der Erhebung empirischer Daten unter Lehrenden und Sozialwissenschaftlern nach wie vor Skepsis.² Neben anderen Kritikpunkten ist ein weit verbreiteter und für die meisten wohl der schwerwiegendste Einwand gegen ihre Verwendung der, dass die erreichbaren Rücklaufquoten zu gering seien, Repräsentativität der Ergebnisse aufgrund massiver Selbstrekrutierung entsprechend nicht sichergestellt werden könne (u. a. Diekmann 1998: 333f., 439ff.; Häder 2006: 237; Atteslander 2008: 147; Schnell/Hill/Esser 2008: 359).

1 Den Herausgebern und zwei anonymen Gutachtern danke ich für wertvolle Kommentare und Anregungen.

2 So finden sich in weithin rezipierten Lehrbüchern der Methoden empirischer Sozialforschung eher abwertende bzw. einseitig kritische Darstellungen der postalischen Befragung (z. B. Friedrichs 1990; Atteslander 2008; Schnell/Hill/Esser 2008). Im Gegensatz dazu kann man bei Häder (2006) sowie Diekmann (1998; 2007) positivere bzw. ausgewogenere Abhandlungen nachlesen. Auch die Anwendung dieser Methode in der Forschungspraxis ist angesichts ihrer vielfältigen Möglichkeiten und den neueren Problemen anderer Erhebungsmodi (etwa sinkende Ausschöpfung in Face-to-Face Befragungen) bislang eher gering: „Den Jahresberichten der ADM-Institute folgend [...] werden von den genannten Instituten in Deutschland gegenwärtig etwa 20 % der quantitativen Interviews postalisch durchgeführt“ (Häder 2006: 234).

Diese Argumentation spiegelt nach Ansicht der Autorin in zweifacher Hinsicht veraltete bzw. uninformierte Auffassungen wider: *Erstens* stellt die Ausschöpfungsquote von Befragungen lediglich *einen* – und nicht einmal einen besonders aussagekräftigen – Indikator für die Güte von Erhebungen dar (Groves 1989: 133, 147; Platek/Särndal 2001; Biemer/Lyberg 2003)³. Und *zweitens* liegen mittlerweile zahlreiche Forschungsergebnisse vor, die für verschiedene Anwendungsbezüge zeigen, dass postalische Befragungen erfolgreich einsetzbar sind und „zufriedenstellende“ Ausschöpfungsquoten, die teilweise sogar jene von renommierten Face-to-Face-Befragungen (z. B. ALLBUS) übertreffen, erzielt werden können (vgl. z. B. Hippler/Seidel 1985; Niedermayer 1989; De Leeuw 1992; Blasius/Reuband 1996; Arzheimer/Klein 1998; Reuband 1999a, 1999b, 2001; Becker/Imhof/Mehlkop 2007; Kunz 2007; Mehlkop/Becker 2007). Sämtliche Publikationen weisen unisono darauf hin, dass qualitativ hochwertige postalische Befragungen (inklusive hoher Ausschöpfungsquoten) v. a. durch die Umsetzung zentraler Empfehlungen der Total bzw. Tailored Design Method (Dillman 1978, 2000; Dillman/Smyth/Christian 2009) gelingen.

Dass mehrmalige Kontaktierungen der angeschriebenen Personen – sog. Mahn- bzw. Nachfassaktionen – für die Steigerung der Rücklaufquote von zentraler Bedeutung sind⁴, ist vielfach belegt (Toops 1926; Yu/Cooper 1983; Guadagnoli/Cunningham 1989; Sobal/Ferentz 1989; Lamnek/Trepl 1991; Yammarino/Skinner/Childers 1991; Blasius/Reuband 1996 u. v. a.).

Doch welchen Einfluss haben Mahnaktionen auf das Antwortverhalten der Respondenten, d. h. auf das Ausmaß an Item-Nonresponse sowie auf die Antworten? Und inwiefern haben mehrfache Kontaktierungen Auswirkungen auf die soziale Zusammensetzung der Befragtenstichprobe? Welche Bedeutung haben Mahnaktionen damit, jenseits ihrer rücklaufsteigernden Wirkung, für die Güte der Befragung? Anders ausgedrückt: Ist es möglich, dass durch Mahnaktionen zwar die Ausschöpfungsquote erhöht, die Repräsentativität der Stichprobe aber verschlechtert wird? Erbringen geringere Ausschöpfungsquoten also u. U. validere Ergebnisse? Nur vergleichsweise wenige Studien beschäftigten sich bislang mit diesen Fragestellungen. Vorliegender Artikel untersucht die benannten Aspekte anhand der Daten einer im Jahr 2009 im Raum Südbaden durchgeführten postalischen Befragung (N=3.555) mit insgesamt drei Erhebungszeitpunkten zum Thema „Kriminalität im höheren Lebensalter“.

3 Einige Autoren konnten anhand empirischer Analysen zeigen, dass höhere Ausschöpfungsquoten nicht zwangsläufig mit verbesserter Stichprobenqualität, d. h. mit geringerem Ausmaß systematischer Antwortausfälle, einhergehen (Groves 1989; Guadagnoli/Cunningham 1989; Sobal/Ferentz 1989; Koch 1998; Schneekloth/Leven 2003; Stoop 2005: 286ff.). Entsprechend umstritten ist auch, welche Ausschöpfungsquoten als „gut“ bzw. „schlecht“ zu bewerten sind (für einen Überblick zu verschiedenen Positionen vgl. Stoop 2005: 23).

4 Daneben kann u. a. auch die Verwendung von Anreizen deutlich zur Rücklaufsteigerung beitragen (vgl. dazu kurz Abschnitt 4.2).

Der folgende Abschnitt führt zunächst einige theoretische Überlegungen zur allgemeinen Wirkungsweise von Nachfassaktionen in postalischen Befragungen an. Anschließend werden der Forschungsstand und die analytische Vorgehensweise der Studie skizziert (Abschnitt 3). Im vierten Abschnitt wird die Befragung, welche die Datenbasis liefert, kurz vorgestellt. Die Darstellung der Studienergebnisse erfolgt in Abschnitt 5. In einem letzten Schritt werden die Befunde zusammengefasst und kritisch reflektiert sowie methodologische Implikationen für die Durchführung bzw. Bewertung postalischer Befragungen benannt (Abschnitt 6).

2 Theoretische Überlegungen zur Wirkungsweise von Nachfassaktionen

Mehrmalige Kontaktierungen der Befragungspopulation werden durchgeführt, um auch Personen, die nicht unmittelbar nach dem Erstversand des Fragebogens antworten, zu einer Survey-Teilnahme zu bewegen. Die Wirkungsweise von Erinnerungsschreiben beruht dabei auf verschiedenen Mechanismen.

1. Ein großer Teil des Nonresponse ist keine stabile, sondern eine situationsabhängige Verhaltensweise (Schnell 1997: 191). Viele Personen sind prinzipiell zu einer Teilnahme bereit, vergessen jedoch das Ausfüllen des Fragebogens – etwa weil der Erstkontakt zu einem ungünstigen Zeitpunkt stattfand – oder schieben die Beantwortung der Fragen hinaus. Diese Personengruppe wird durch das Nachfassen schlicht an den Erhalt des Fragebogens erinnert. Hierdurch dürfte die Teilnahmerate steigen.

Aus Sicht der sozialen Austauschtheorie (Homans 1961; Blau 1964), die v. a. von Dillman (1978, 2000; Dillman/Smyth/Christian 2009) im Rahmen seiner Total/Tailored Design Method zur Erklärung des Teilnahmeverhaltens in postalischen Befragungen herangezogen wird, sind Reminder eine (weitere) Maßnahme des Forschers mit dem Ziel, das Nutzenkalkül der angeschriebenen Personen in Bezug auf die Befragungsteilnahme positiv zu beeinflussen. Damit dies gelingt, sollte durch die erneute Kontaktierung der subjektive Nutzen der Respondenten gesteigert und/oder deren Kosten gesenkt und/oder deren Vertrauen gesteigert werden. Es kommen entsprechend zwei weitere Wirkungsmechanismen in Betracht:

2. Zwar bleiben durch Follow-ups die Kosten einer Befragungsteilnahme konstant. Der Nutzen jedoch wird gesteigert, indem deutlich gemacht wird, dass die Meinung und Unterstützung des Angeschriebenen für den Erfolg der Untersuchung von zentraler Bedeutung sind. Zudem unterstreicht die Erinnerungsaktion die Sinnhaftigkeit der Studie und die Ernsthaftigkeit der Forscherabsichten. Dadurch kann das Vertrauen der Befragten gewonnen bzw. stabilisiert werden (Petermann 2005: 62f.).

3. In vorliegender Untersuchung wurde allen Angeschriebenen bereits beim Erstkontakt eine Fünf-Euro-Banknote bedingungslos und gerahmt als „kleines Dankeschön“ beigelegt (Näheres siehe Abschnitt 4.2). In Wechselwirkung mit der Vergabe eines initialen (monetären) Anreizes erscheint nun noch eine weitere Funktionsweise von Remindern plausibel, die im Vergleich zur vorher skizzierten eine stärkere normative Konnotation besitzt. Follow-ups erinnern an den Erhalt des Incentives und verstärken damit das Gefühl der sozialen Verpflichtung (Reziprozitätsnorm). Der Anreiz für eine – wenn auch verspätete – Befragungsteilnahme besteht nun v. a. in der Vermeidung eines schlechten Gewissens.

3 Stand der Forschung und Vorgehensweise der Studie

Zunächst werden bisherige empirische Befunde zu den Auswirkungen von Mahnaktionen auf den Item-Nonresponse, die soziale Zusammensetzung der Stichprobe und die Teilnahmebereitschaft von Menschen mit verschiedenen Einstellungen, Eigenschaften, Erfahrungen und Verhaltensweisen überblicksartig vorgestellt.

Auswirkungen auf das Ausmaß an Item-Nonresponse (INR): Hippler/Seidel (1985) und Reuband (1999b, 2001) konstatieren einen kontinuierlichen Anstieg fehlender Antworten mit jeder Nachfassaktion. So steigerte sich etwa bei Reuband (1999b: 95) der durchschnittliche Anteil fehlender Antworten an allen gestellten Fragen von 4,8 % beim Erstkontakt auf 7,2 % bei der vierten Kontaktierung.

Auswirkungen auf die soziale Zusammensetzung der Stichprobe: Während einige Forscher keine Unterschiede zwischen Respondenten, die früher oder später antworten konstatieren (Koenig/Martin/Seiler 1977; Goudy 1978; Guadagnoli/Cunningham 1989; Sobal/Ferentz 1989), stellen andere Studien fest, dass das Alter der Befragten mit den Nachfassungen variiert. Der konsistente Befund lautet hier, dass Menschen ab einem Alter von 60/65 Jahren überwiegend bereits beim Erstkontakt antworten, während mit weiteren Nachfassaktionen v. a. jüngere Personen in die Stichprobe rekrutiert werden (Hippler/Seidel 1985: 50; Blasius/Reuband 1996; Reuband 1999b: 88, 2001: 320). Scott (1961), Brambilla/McKinlay (1987), Blasius/Reuband (1996), Reuband (1999b; 2001) sowie Petermann (2005) konnten ferner zeigen, dass Angehörige unterer sozialer Schichten, d. h. Personen mit geringer Bildung sowie wenig qualifizierten Berufen, erst mithilfe von Nachfassaktionen stärker repräsentiert werden. Allerdings konnten *keine* Unterschiede im Nettoeinkommen der Befragten verschiedener Erhebungszeitpunkte festgestellt werden (Reuband 1999b; Petermann 2005: 72f.). Die einzige Studie, die Hinweise darauf fand, dass Mahnaktionen die Stichprobenanteile der Geschlechter beeinflussen, stammt von Niedermayer (1989). Hier antworteten Frauen später als Männer.

Aus der Surveyforschung ist bekannt, dass die Thematik einer Befragung bzw. deren Bewertung durch die Respondenten einen Einfluss auf die Teilnahmebereitschaft der Befragten hat (Kanuk/Berenson 1975: 449; Dommeyer 1985; Hippler/Seidel 1985; Schnell 1997: 170ff.), und dass in Populationen, deren Merkmale eng mit dem Befragungsthema zusammenhängen (sog. Spezialpopulationen), zumeist höhere Rücklaufquoten erzielt werden als in allgemeinen Bevölkerungsstichproben (Hippler 1988). Bei genauerer Betrachtung der hier angeführten Studien fällt nun auf, dass die jeweils untersuchten Populationen und Themen ebenfalls einen Einfluss auf die Wirkungsweise von Nachfassaktionen haben. Für Spezialpopulationen im oben genannten Sinne können im Vergleich zu allgemeinen Bevölkerungsstichproben viel seltener Unterschiede zwischen den Respondenten verschiedener Erhebungszeitpunkte festgestellt werden. Nachfassaktionen erhöhen hier lediglich den Rücklauf, haben jedoch kaum Auswirkungen auf die Zusammensetzung der Netto-Stichprobe. Für allgemeine Bevölkerungsstichproben bzw. Populationen, deren Merkmale nicht mit dem Befragungsthema korrelieren, tragen Reminder hingegen nicht nur zur Erhöhung der Ausschöpfungsquote, sondern auch zu einer Veränderung (bzw. Verbesserung) der Stichprobenszusammensetzung bei.⁵

Insgesamt kann festgehalten werden, dass sich die Befragten der einzelnen Mahnaktionen zwar hinsichtlich verschiedener sozio-demografischer Merkmale teilweise deutlich voneinander unterscheiden. Bisherige Untersuchungen deuten aber darauf hin, dass sich die sozialen Zusammensetzungen der gesamten Befragtenstichproben *vor* und *nach* der Durchführung von Mahnaktionen in Bezug auf die Merkmale Alter, Geschlecht und Bildung sehr ähnlich sind. „Die Prozentpunktunterschiede sind nicht allzu groß – und dies gilt selbst in gewissem Maße für die Variablen, für die nennenswerte Effekte der Mahnaktion festgestellt werden können“ (Reuband 2001: 328).

Dennoch: während Reuband (2001: 328) zu der Feststellung gelangt, dass „[...] Mahnaktionen zur Verbesserung der sozialen Zusammensetzung [...] beitragen und deshalb prinzipiell sinnvoll sind“, kommen etwa Koch (1998) sowie Schneekloth/Leven (2003) anhand ihrer Analysen verschiedener ALLBUS-Erhebungen⁶ zu dem Schluss, dass die ohnehin nur geringen Abweichungen sozialer Merkmale der Befragten von denen der Grundgesamtheit durch Nachbearbeitungen weder abgeschwächt noch beseitigt werden. Es erhöht sich vorrangig die Fallzahl, nicht die Qualität der Stichprobe. Petermann (2005) behauptet gar, dass bestehende Ver-

5 Zu den gleichen Ergebnissen gelangt Leslie (1972) in seiner Metastudie.

6 Zwar wird der ALLBUS mittels persönlicher Interviews erhoben und stellt somit keine postalische Befragung dar. Dennoch weisen die Forschungsfragen bzw. Erkenntnisse beider Studien mode-unabhängig einen engen Bezug zu den im vorliegenden Beitrag behandelten Fragestellungen auf.

zerrungen in postalischen Befragungen durch Nachfassungen (zumindest minimal) verstärkt werden und deshalb auf Mahnaktionen sogar verzichtet werden sollte.⁷

Angesichts der einhelligen⁸ Feststellung, dass sich die sozialen Merkmale der Befragten vor und nach Nachfassungen nur geringfügig unterscheiden und zudem kaum von entsprechenden Verteilungen in der Grundgesamtheit abweichen,⁹ schlussfolgern sämtliche benannten Autoren übereinstimmend, dass auch niedrige Ausschöpfungsquoten eine gute Repräsentativität der Stichprobe erbringen können.

Auswirkungen auf die Teilnahmebereitschaft von Menschen mit verschiedenen Einstellungen, Eigenschaften, Erfahrungen und Verhaltensweisen: Zahlreiche weitere Merkmale, die in keinem oder nur sehr schwachem Zusammenhang mit den geprüften Sozio-Demografika stehen, sind für die Charakterisierung der Grundgesamtheit bzw. der Stichprobe von Bedeutung. Deshalb erbringt ein Vergleich verschiedener Merkmale der Befragten einzelner Erhebungswellen Erkenntnisse darüber, ob Mahnaktionen zu Veränderungen des Samples führen.¹⁰

Reuband (1999b, 2001) und Petermann (2005) konnten feststellen, dass mit zunehmenden Mahnaktionen der Anteil politisch Uninteressierter steigt, während thematisch interessierte Personen vorwiegend zu Beginn der Erhebung teilnehmen.

Keine Auswirkungen der Mahnaktionen fand Reuband (1999b: 93ff.) hinsichtlich des Vertrauens der Befragten in ihre Mitmenschen, in Bezug auf ihre Lebenszufriedenheit, die Anzahl möglicher Sorgen, hinsichtlich der Kriminalitätsfurcht, eigener Opfererfahrungen sowie in Bezug auf die politische Orientierung.

Zusammenfassend kommen die Studien zu dem Schluss, dass sich die Befragten nach dem Erstversand auch hinsichtlich „weicher“ Merkmale kaum von denen der Gesamtheit nach allen Mahnaktionen unterscheiden. Die Effekte der Reminder sind schwach bzw. nicht signifikant (Reuband 1999b: 96). Einstellungen und psychosoziale Befindlichkeiten der Respondenten werden durch Mahnaktionen reproduziert.

Vorgehensweise der Studie: Zwar ist die rücklaufsteigernde Wirkung von Nachfassungen bereits umfangreich belegt. Der Vollständigkeit halber soll dies dennoch anhand der Südbadener postalischen Befragung kurz veranschaulicht und kom-

7 Zu ähnlichen Ergebnissen gelangt Stoop (2005: 287) in Bezug auf den Nonresponse im European Social Survey (ESS): auch hier wurde die Stichprobenqualität durch einseitige Konvertierung anfänglicher Verweigerer eher verschlechtert.

8 Ausnahme ist bei Reuband (1999b: 91) das Bildungsniveau; dieses ist in der Stichprobe im Vergleich zur Population (wie häufig in postalischen Befragungen) leicht erhöht.

9 Es sei an dieser Stelle noch einmal darauf hingewiesen, dass derartige „Repräsentanzabgleiche“ lediglich als Indiz, keinesfalls als Beweis für eine gelungene Abbildung der Grundgesamtheit durch die Stichprobe gelten können (vgl. zum „Repräsentanzbegriff“ auch Schnell/Hill/Esser 1995: 287).

10 Da ein Abgleich von Einstellungen, persönlichen Erfahrungen, Verhaltensweisen der Befragten etc. mit offiziellen Daten nicht möglich ist, kann allerdings nicht festgestellt werden, welche Auswirkungen dies auf eventuell bereits beim Erstkontakt bestehende Verzerrungen hat.

mentiert werden. Anschließend werden die Auswirkungen von Mahnaktionen auf a) das Ausmaß an Item-Nonresponse, b) die soziale Zusammensetzung der Stichprobe sowie auf c) die Teilnahmebereitschaft von Menschen mit verschiedenen Einstellungen, Eigenschaften, Erfahrungen und Verhaltensweisen untersucht. Die Ergebnisse werden den bisherigen Forschungsbefunden jeweils gegenübergestellt. Um die Auswirkungen von Mahnaktionen auf die benannten Aspekte analysieren zu können, wurden die eingehenden Fragebögen anhand ihres Eingangsstempels den einzelnen Erhebungswellen zugeordnet.¹¹

Vor Beginn der eigentlichen Analyse in Abschnitt 5 wird jedoch zunächst die verwendete Datenbasis vorgestellt.

4 Die Datengrundlage

Die Fragestellungen des vorliegenden Beitrages werden auf der Grundlage der Daten einer regionalen postalischen Befragung mit dem Titel „Alltag & Gesetze. Erfahrungen und Einstellungen in der zweiten Lebenshälfte“ untersucht. Diese Befragung wurde im Jahr 2009 im Rahmen eines Forschungsprojekts zum Thema „Ältere Menschen und Kriminalität“ am Max-Planck-Institut für ausländisches und internationales Strafrecht in Freiburg durchgeführt.¹²

4.1 Grundgesamtheit, Stichprobe und Fragebogen

Zur Grundgesamtheit der Befragung zählen deutsche Staatsbürger, die zwischen dem 01.01.1929 und dem 01.01.1959 geboren wurden und in Privathaushalten der 302 Gemeinden des Regierungsbezirkes (RB) Freiburg, d. h. in der Region Südbaden, wohnhaft sind.¹³ Die Bruttostichprobe beruht auf einer geschichteten, geklumpten, zweistufigen Zufallsauswahl und enthält 3.555 Personen im Alter zwischen 49 und 81 Jahren, die aus den Adressregistern von Einwohnermeldeämtern zufällig gezogen wurden.

11 Da die Abstände zwischen den einzelnen Erhebungswellen recht groß waren, dürften die Unschärfen, die sich aus dieser Kodierung ergeben, geringfügig sein.

12 Projektbeschreibung: <http://www.mpicc.de/ww/de/pub/forschung/forschungsarbeit/kriminologie/aelteremenschen.htm> (07.06.2010).

13 Menschen, die in Institutionen wie etwa Alters- oder Pflegeheimen leben, sowie Personen mit ausländischer Staatsbürgerschaft sind nicht in der Population enthalten; für diese können im Rahmen der Studie entsprechend keine Aussagen getroffen werden.

Der verwendete Fragebogen umfasst 24 Seiten und beinhaltet 85 Fragen bzw. 240 einzelne Items.¹⁴ Er wurde als DIN A4-Broschüre auf farbiges Papier gedruckt. Der Bogen beinhaltet Fragen zu Einstellungen gegenüber Gesetzen und Gesetzesübertretungen, zu Viktimisierungserfahrungen, zu Kriminalitätsfurcht sowie zu bisheriger polizeilicher Registrierung, kriminellen Verhalten innerhalb verschiedener Lebensabschnitte und zur Delinquenzbereitschaft. Daneben wurden Informationen zu Persönlichkeitseigenschaften, Werteinstellungen, bedeutenden Lebensereignissen sowie zum emotionalen Wohlbefinden, dem Gesundheitszustand und zu Einstellungen gegenüber dem Altern erhoben. Sozio-demografische Angaben wurden am Ende des Bogens abgefragt.¹⁵ Aufgrund der sensitiven Forschungsthematik „selbstberichtete Delinquenz“ wurde die Befragung vollständig anonym, d. h. ohne Aufdruck einer Identifikationsnummer, durchgeführt.¹⁶ Um dennoch effizient nachfassen zu können, wurden adressierte und frankierte Antwortpostkarten beigelegt, die nach erfolgter Beantwortung des Fragebogens – quasi als Benachrichtigung – seitens der Respondenten mit Namen versehen separat an uns zurückgeschickt werden sollten.

4.2 Datenerhebung, Ausschöpfung und weitere methodenbezogene Befunde

Die Erhebungsphase erstreckte sich von Juni bis August 2009, betrug ca. neun Wochen und beinhaltete drei Erhebungszeitpunkte.¹⁷ Beim Erstkontakt wurde an alle Befragten ein DIN C4-Umschlag verschickt, welcher ein doppelseitiges Anschreiben, den Fragebogen, einen adressierten und frankierten Rückumschlag sowie eine adressierte und frankierte Antwortpostkarte enthielt. Als forschungsdurchführende Institutionen wurden das Max-Planck-Institut Freiburg sowie das Institut für Soziologie der Universität Freiburg genannt. Zur Steigerung des Rücklaufs wurde allen Ange-

14 Davon sollten 76 Fragen bzw. 164 Items von allen Befragten beantwortet werden, während weitere 9 Fragen bzw. 76 Items als sog. Filterfragen nur in Abhängigkeit vorheriger Antworten gegebenenfalls zu beantworten waren.

15 Das Erhebungsinstrument wurde vor dem Einsatz mittels „Zwei-Phasen-Pretesting“ (Prüfer/Rexroth 2000), welches einen kognitiven Pretest (n=8) und einen Standardpretest (N=400) beinhaltete, schrittweise geprüft und optimiert.

16 Die Ergebnisse eines im Vorfeld durchgeführten Methodenexperiments zu den Auswirkungen anonymer bzw. nicht-anonymer Befragung auf die Rücklaufquote, das Antwortverhalten und die Stichprobenszusammensetzung trugen ebenfalls zu dieser Entscheidung bei.

17 Die Konzeption der postalischen Befragung orientiert sich im Wesentlichen an der Total/Tailored Design Method (Dillman 1978, 2000; Dillman/Smyth/Christian 2009). Die Anzahl, die Zeitpunkte sowie die Form der einzelnen Mahnaktionen weichen allerdings etwas von den Empfehlungen ab.

schriebenen bereits beim Erstkontakt eine Fünf-Euro-Banknote beigelegt.¹⁸ Diese Gabe wurde im Anschreiben mit folgendem Satz „gerahmt“: „*Als kleine Aufwandsentschädigung und Dankeschön für Ihre Mitwirkung haben wir Ihnen 5 Euro beigelegt. Diese können Sie in jedem Fall behalten.*“ Um die Angeschriebenen darüber hinaus zur Unterstützung des Forschungsvorhabens zu motivieren, wurde dieses in einer regionalen Sonntagszeitung, in den jeweiligen Amts-/Gemeindeblättern der gesampelten Ortschaften sowie im Hörfunk vorgestellt.

Etwa zweieinhalb Wochen nach dem Erstversand wurde ein kurzes Erinnerungsschreiben an diejenigen Personen verschickt, von denen bisher keine Antwortkarte mit dem Hinweis bei uns eingegangen war, dass der Fragebogen ausgefüllt retourniert wurde. Dieses Schreiben enthielt sowohl die erneute Bitte um Teilnahme als auch – für den Fall bereits erfolgter Rücksendung des Fragebogens – Dank für die geleistete Unterstützung. Nach weiteren drei Wochen wurden die Befragten ein letztes Mal kontaktiert, indem noch einmal sämtliche Unterlagen des Erstversands sowie ein letztes Erinnerungsschreiben zugeschickt wurden.

Von den 3.555 angeschriebenen Personen fielen 132 neutral aus,¹⁹ 1.997 Fragebögen konnten in die Auswertung einbezogen werden. Auf der Basis der bereinigten Bruttostichprobe erzielte die Befragung somit eine Ausschöpfungsquote²⁰ in Höhe von 58,4 %.

Insgesamt wurden 9 % aller vergebenen Incentives zurückerhalten. Die beigelegte Fünf-Euro-Banknote wurde dabei sowohl von Personen, die aus verschiedenen Gründen nicht an der Befragung teilnehmen konnten (neutrale Ausfälle), von Personen, die nicht an der Befragung teilnehmen wollten (Verweigerer), als auch von Angeschriebenen, die den Fragebogen beantwortet hatten, zurückgegeben.

Die *separate-Postkarten-Technik* (sPT), die zur Rücklaufkontrolle eingesetzt wurde, funktionierte in Anbetracht der zusätzlichen Komplexität der Aufgabenstellung und des Alters der Befragten recht gut. Insgesamt teilten uns 1.431 Personen mit, dass sie einen Fragebogen retourniert haben. Geht man davon aus, dass

18 Der Einsatz von Anreizen bei der Durchführung postalischer Befragungen wird u. a. von Dillman (2000: 150) und Dillman/Smyth/Christian (2009: 238ff.) im Rahmen der Tailored Design Method empfohlen. Zur rücklaufsteigernden Wirkung monetärer Incentives speziell in postalischen Befragung zu kriminalitätsbezogenen Themen vgl. Becker/Imhof/Mehlkop (2007); Mehlokop/Becker (2007). Dort werden auch zahlreiche weitere Studien zum Einsatz monetärer Anreize in postalischen Befragungen benannt.

19 Als neutrale Ausfälle gelten: unbekannt/verzogen, verstorben, Krankheit/Pflege/Behinderung, kaum deutsche Sprachkenntnisse, Person gehört nicht zur Grundgesamtheit. Die bereinigte Bruttostichprobe umfasst somit 3.423 Personen.

20 Die Ausschöpfungsquote (A) wurde hier wie folgt berechnet (Schnell/Hill/Esser 1995: 289; Diekmann 1998: 359): $A = (\text{Anzahl auswertbarer Fragebögen} / \text{Umfang der bereinigten Bruttostichprobe}) * 100 \%$.

dies lediglich Personen waren, die tatsächlich an der Befragung teilgenommen haben, dann entspricht dies einer beachtlich guten Benachrichtigungsquote²¹ von 71,7 %. Die Ergebnisse zeigen zweierlei: 1) die Anzahl der retournierten Postkarten ist *nicht* – wie im Lehrbuch von Schnell/Hill/Esser (2008: 363) nach wie vor behauptet – höher als die der retournierten Fragebögen (vgl. ähnliche Ergebnisse bei Reuband 1999a) und 2) die sPT bietet eine gute Möglichkeit, *sowohl* Anonymität *als auch* effektive Rücklaufkontrolle in postalischen Befragungen zu kombinieren – und dies sogar mit der Zielgruppe älterer Menschen.

Ein häufig genannter Kritikpunkt gegenüber postalischen Befragungen ist, dass die Erhebungssituation nicht kontrolliert werden kann, es also z. B. nicht sicher ist, wer den Fragebogen ausgefüllt hat (z. B. Diekmann 1998: 439; Schnell/Hill/Esser 2008: 359). Um zumindest das Ausmaß gemeinschaftlicher Beantwortung abschätzen zu können, wurden die Respondenten am Ende des Bogens gefragt: „*Haben Sie den Fragebogen allein oder gemeinsam mit einer anderen Person ausgefüllt?*“ 7,9 % der Befragten gaben an, den Fragebogen gemeinsam mit einer anderen Person ausgefüllt zu haben, die restlichen 92,1 % beantworteten die Fragen allein (n=1.962). Geht man von der Wahrhaftigkeit dieser Angaben aus, sind diese Werte sehr befriedigend. Wie später noch gezeigt wird, nimmt der Anteil derjenigen, die den Fragebogen gemeinschaftlich ausfüllen, allerdings mit jeder Mahnaktion zu (vgl. Abschnitt 5.1).

5 Ergebnisse

Wie bereits in Abschnitt 3 angedeutet, werden die unterschiedlichen Auswirkungen von Mahnaktionen jeweils separat vorgestellt: zunächst wird der Einfluss auf den Gesamtrücklauf betrachtet sowie untersucht, inwiefern sich das Ausmaß an Item-Nonresponse in Abhängigkeit der Kontaktierungen verändert (Abschnitt 5.1). Anschließend werden die Auswirkungen von Nachfassaktionen auf die Stichprobenszusammensetzung und auf verschiedene weitere Merkmale der Befragten behandelt (Abschnitte 5.2 und 5.3).

21 Diese Quote liegt deutlich über bisher dokumentierten Erfahrungen, etwa den von Reuband (1999: 58) berichteten Raten für allgemeine Bevölkerungsbefragungen in Dresden (66 %) und Düsseldorf (60 %).

5.1 Auswirkungen auf den Gesamtrücklauf und den Item-Nonresponse

Die erste Zeile von Tabelle 1 zeigt erneut, was bereits durch zahlreiche Studien belegt wurde: Nachfassaktionen, wie in der Total/Tailored Design Method (Dillman 1978, 2000; Dillman/Smyth/Christian 2009) vorgeschlagen, sind eine erfolgreiche Maßnahme zur Rücklauferhöhung. Allerdings erscheint eine dritte oder gar vierte Nachfassung entgegen Dillman'scher Empfehlungen angesichts der deutlich sinkenden Beiträge der einzelnen Kontaktierungen zum Gesamtrücklauf kaum mehr rentabel (vgl. auch Blasius/Reuband 1996; Reuband 1999b: 86, 2001: 319). Die Netto-Ausschöpfungsquoten der einzelnen Erhebungswellen sind im Gegensatz zu der von Dillman (1978: 185f.) bzw. Hippler (1988: 245) proklamierten „Faustregel“ *nicht* in etwa gleich hoch, sondern nehmen linear ab (vgl. Tab. 1).

Tabelle 1 Auswirkungen von Maßnahmen zur Rücklauferhöhung

	Erstversand	1. Mahnung	2. Mahnung		
1) Gesamtrücklaufquote nach ...	33,4	50,1	58,3		
2) Netto-Ausschöpfungsquote	33,4 (ber. N=3423)	24,9 (ber. N=2286)	16,3 (ber. N=1734)		
	Erstversand (n=1144)	1. Mahnung (n=570)	2. Mahnung (n=283)	Gesamt (n=1997)	Spearman Rho
3) Item-Nonresponse-Rate gesamter Fragebogen	6,13	7,80	7,50	6,80	0,051*
4) Item-Nonresponse-Rate 14 Prävalenzen, Lebensspanne	1,59	3,32	2,60	2,23	0,048*
5) Item-Nonresponse-Rate 14 Inzidenzen, seit 50. Geburtstag	2,37	4,05	3,81	3,05	0,017
6) Item-Nonresponse-Rate 14 Inzidenzen, letzte 12 Monate	4,00	5,10	4,42	4,37	-0,014

Gesamtrücklaufquote nach Erhebungswelle (ber. N=3423; in Prozent), Netto-Ausschöpfungsquoten der Erhebungswellen (in Prozent), Item-Nonresponse-Raten nach Erhebungswellen (in Prozent; grau: die Item-Nonresponse-Rate des Erstversands ist sign. geringer als die der 1. und 2. Mahnung, $p \leq 0,05$).

* $p \leq 0,05$

zu 3) $F=7,734$; $p(2\text{-seit.})=0,00$; $df=2$; zu 4) $F=3,438$; $p(2\text{-seit.})=0,03$; $df=2$; zu 5) $F=4,026$; $p(2\text{-seit.})=0,02$; $df=2$; zu 6) $F=1,209$; $p(2\text{-seit.})=0,30$; $df=2$

Was das Ausmaß an Item-Nonresponse betrifft, so beträgt der durchschnittliche Anteil fehlender Antworten am gesamten Fragebogen in der Südbadener Befragung 6,8 %. Im Mittel fehlen damit ca. 14 Antworten (von 202) pro Fragebogen. Dies stellt einen unauffälligen Wert dar.²² Bemerkenswert ist, dass die durchschnitt-

22 Nach Hippler (1988) liegt die INR-Rate postalischer Befragungen üblicherweise zwischen 1 und 8 %. Das Ausmaß fehlender Antworten in der Südbadener Befragung liegt mit einer INR-Rate von 6,8 % also am oberen Ende des Durchschnitts. Es wird vermutet, dass das Alter der Befragten hierfür verantwortlich sein könnte, da zwischen dem Ausmaß fehlender Antworten und dem Alter der Befragten ein Zusammenhang in Höhe von $r=0,224$ ($p \leq 0,01$) besteht.

lichen INR-Raten sämtlicher Fragen zur Delinquenz der Befragten (auf Itemebene) und zum Netto-Einkommen lediglich zwischen 2 und 4,4 % betragen. Bei den heiklen Fragen sind im Vergleich zur gesamten Befragung also unterdurchschnittlich wenige Antwortausfälle zu verzeichnen. Tabelle 1 gibt für verschiedene Fragenkomplexe einen Überblick über das Ausmaß der Antwortausfälle in Abhängigkeit der einzelnen Erhebungswellen. Mittels einfaktorieller Varianzanalyse wurde jeweils geprüft, ob sich die Item-Nonresponse-Raten in Abhängigkeit des Erhebungszeitpunktes unterscheiden. Anhand des Student-Newman-Keuls-Tests wurde zudem ermittelt, *welche* Erhebungswellen sich signifikant voneinander unterscheiden.

Aus Tabelle 1 geht deutlich hervor, dass das Ausmaß fehlender Antworten mit den Mahnaktionen ansteigt. Während jedoch Hippler/Seidel (1985) und Reuband (1999b, 2001) einen kontinuierlichen Anstieg fehlender Antworten über die Mahnaktionen feststellen, zeigt sich hier ein kurvilinearere Zusammenhang. Sämtliche INR-Raten erreichen ihren Höhepunkt unter den Respondenten der 1. Mahnung, sinken dann leicht wieder ab, bleiben aber über dem Niveau der nach dem Erstversand retournierten Fragebögen.²³

Reuband (2001: 321f.) konnte zeigen, dass es durch die Mahnaktionen gelingt, auch Personen mit Zeitmangel und tendenziell nicht-kooperative Personen in die Befragung einzubeziehen. Dadurch erklärt sich seines Erachtens auch der unter den nachträglich rekrutierten Respondenten erhöhte INR (2001: 322). Mit den Worten Reubands (2001: 322) kann deshalb Folgendes festgehalten werden:

„Man mag dies [Anm. d. Autorin: den INR-Anstieg] einerseits als Nachteil für die Qualität der Daten ansehen: gilt doch, dass diese Personen weniger motiviert sind und den Fragebogen weniger sorgfältig beantworten. Andererseits bedeutet es auch, dass Personen mit wenig auskristallisierten Einstellungen und 'non-attitudes' (Converse 1964) erfasst werden. Widersprüche in den Einstellungen und Meinungslosigkeit sind Bestandteil der sozialen Realität und bedürfen der Berücksichtigung.“

Zudem erscheint der Trade-off zwischen verbesserter Stichprobenqualität und verschlechterter Datenqualität – sofern man Item-Nonresponse überhaupt als Beeinträchtigung der Datenqualität wertet – in Anbetracht der zwar meist signifikanten,

23 Eine mögliche Erklärung für die INR-Verteilung könnte sein, dass Personen ohne Schulabschluss bzw. mit Volks-/Hauptschulabschluss im Vergleich zum Erstversand und zur 2. Mahnaktion am stärksten auf die 1. Mahnung reagiert haben, während Personen mit höheren Bildungsabschlüssen anteilmäßig sowohl im Erstversand als auch in der 2. Mahnung stärker vertreten sind als in der 1. Mahnung. (Diese Ergebnisse sind hier nicht explizit dargestellt, können aber auf Anfrage zur Verfügung gestellt werden.) Aufgrund etwas geringerer kognitiver Kompetenzen bzw. weil längeres Lesen und Schreiben ungewohnt ist, könnten deshalb vermehrt Fehler beim Ausfüllen des Fragebogens aufgetreten sein. Eine alternative Erklärung wäre, dass Teilnehmer der 1. Mahnung eher „genervt“ und deshalb unkonzentrierter antworten als Befragte der 2. Mahnaktion, die eventuell aufgrund schlechten Gewissens der sehr späten Teilnahme wiederum etwas sorgfältiger antworten.

aber insgesamt recht geringen Steigerung der Antwortausfälle über die Nachfassaktionen durchaus akzeptabel.

Das Ausmaß fehlender Antworten kann als *ein* Kriterium der Qualität der Erhebungsdaten angesehen werden. Ein weiteres ist die gemeinschaftliche Beantwortung des Fragebogens. Wie bereits weiter oben erwähnt wurde, steigt mit den Mahnaktionen auch der Anteil derjenigen an, die den Fragebogen gemeinsam mit einer anderen Person, statt wie gewünscht allein, ausgefüllt haben: von 5,5 % im Erstversand auf 12,7 % in der 2. Mahnung. Dieser Zusammenhang zwischen gemeinschaftlicher Beantwortung und Beantwortungszeitpunkt ist signifikant (Chi-Quadrat=13,429; df=2; Cramers V=0,083; p(2-seitig)=0,001).

5.2 Auswirkungen auf die soziale Zusammensetzung der Stichprobe

Um zu überprüfen, inwiefern Mahnaktionen die soziale Zusammensetzung der Befragtenstichprobe – und damit ihre Güte – beeinflussen, wurden zunächst die sozio-demografischen Merkmale Geschlecht, Alter (gruppiert) und Wohnort der Befragten (kategorisiert nach Einwohneranzahl) den entsprechenden Verteilungen in der Grundgesamtheit gegenüber gestellt.²⁴ Um Veränderungen der Stichprobenkomposition bezüglich der einzelnen Merkmale zu erkennen, wurden zudem die Merkmalsverteilungen der über die einzelnen Erhebungswellen *akkumulierten* Netto-Stichproben betrachtet. Insgesamt konnte Folgendes festgestellt werden:

- a) Die gesamte Befragtenstichprobe bildet die Grundgesamtheit hinsichtlich des Alters, des Geschlechts und des Wohnorts der Personen recht gut ab. Abweichungen von den Verteilungen der Population liegen jeweils im minimalen Bereich.²⁵
- b) Die Befragten der verschiedenen Erhebungswellen unterscheiden sich hinsichtlich der Geschlechtszugehörigkeit sowie des Alters *nicht* signifikant voneinander. Tendenziell gilt jedoch auch hier, dass ältere Zielpersonen eher antworten als jüngere (vgl. Reuband 1999b; vgl. Reuband 2001; Petermann 2005). Die Art des Wohnortes der Befragten allerdings unterscheidet sich signifikant

24 Die Daten der Grundgesamtheit wurden der Bevölkerungsstatistik 2005 des Statistischen Landesamtes Baden-Württemberg entnommen. Da die Befragung anonym durchgeführt wurde, können sämtliche Abgleiche lediglich auf der Aggregatdatenebene vollzogen werden.

25 Tendenziell beteiligten sich Männer minimal häufiger an der Befragung als Frauen. Dies könnte daran liegen, dass Männer einen stärkeren Bezug zum Thema „Kriminalität“ haben. Personen im Alter zwischen 60 und 70 Jahren sind in der Südbadener Befragung etwas überrepräsentiert, während Angehörige der höchsten Altersgruppe leicht unterrepräsentiert sind. Die Bewohner von Groß- bzw. Mittelstädten sind in der Netto-Stichprobe der gesamten Befragung leicht stärker vertreten als in der Grundgesamtheit, während Dorfbewohner etwas unterrepräsentiert sind.

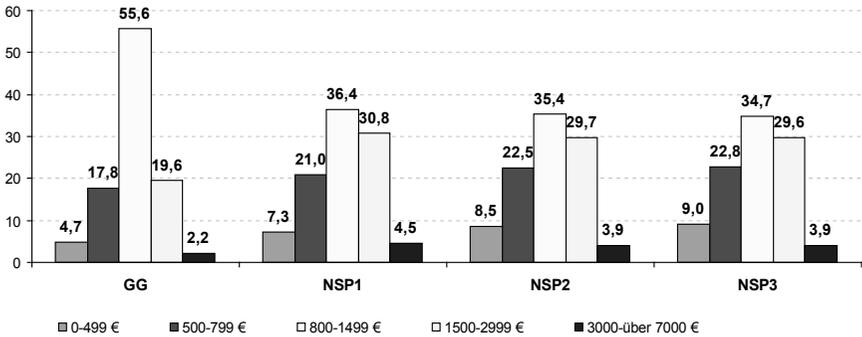
zwischen den Erhebungswellen: der Anteil der Dorfbewohner steigt mit zunehmender Zahl der Kontaktierungen an (Chi-Quadrat=13,050; df=4; Cramers V=0,059; p(2-seitig)=0,011).

- c) Zwar unterscheiden sich die Netto-Stichproben der einzelnen Erhebungswellen nicht wesentlich voneinander (ähnlich Blasius/Reuband 1996; Reuband 1999b, 2001; ähnlich Petermann 2005). Dennoch ist die *gesamte* Befragtenstichprobe im Vergleich zu den Stichproben nach Erstversand bzw. nach Erstversand und 1. Mahnung hinsichtlich der geprüften Merkmale den Verteilungen in der Grundgesamtheit jeweils am ähnlichsten. Dies bedeutet, dass die Durchführung von zwei Mahnaktionen die soziale Komposition der Stichprobe im Sinne einer Annäherung an die Verhältnisse in der Population positiv beeinflusst. Andererseits bedeutet es eben auch, dass bereits die Stichprobe nach dem Erstversand zumindest die Verteilung zentraler sozio-demografischer Merkmale in der Grundgesamtheit relativ gut repräsentiert.

Um weitere Hinweise auf eventuelle Verzerrungen der Stichprobe infolge von Mahnaktionen zu erhalten, wurden ferner das monatliche Personen-Nettoeinkommen, die Schulbildung sowie die Haushaltsgröße der Befragten mit offiziellen Daten der Grundgesamtheit – hier: Daten für den Regierungsbezirk Freiburg des Mikrozensus 2006²⁶ – verglichen. Aus Abbildung 1 geht hervor, dass sowohl Personen der untersten beiden Einkommenskategorien (0 bis 799 Euro) als auch der beiden oberen (1.500 bis über 7.000 Euro) in der gesamten Befragtenstichprobe (NSP3) im Vergleich zur Grundgesamtheit überrepräsentiert sind, während Personen mit einem mittleren Einkommen in Höhe von 800 bis 1.499 Euro deutlich unterrepräsentiert sind. Die Einkommensverteilung der Respondenten der einzelnen Erhebungswellen lässt die Tendenz erkennen, dass die Anteile derjenigen Personen, deren monatliches Netto-Einkommen in den unteren Kategorien liegt, mit zunehmenden Mahnaktionen etwas zunimmt, während der Anteil derjenigen mit Einkommen in den höheren Kategorien leicht abnimmt. Allerdings sind diese Unterschiede zwischen den Erhebungswellen minimal und *nicht* signifikant (ähnliche Ergebnisse vgl. Reuband 2001: 89f.; Petermann 2005: 72f.).

26 Nähere Informationen zum Mikrozensus: <http://www.forschungsdatenzentrum.de/bestand/Mikrozensus/index.asp> (07.06.2010). Da für die Nettovalidierung tief regionalisierte Daten (Regierungsbezirksebene) des Mikrozensus benötigt wurden und diese nicht im Scientific Use File enthalten sind, wurden die Analysen mittels sog. kontrollierter Datenfernverarbeitung durchgeführt.

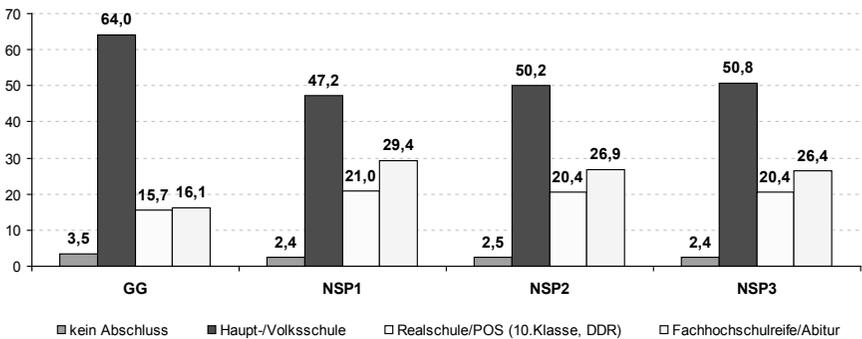
Abbildung 1 Rücklaufquoten nach Personen-Nettoeinkommen



Monatliches Personen-Nettoeinkommen (kategor.) in Grundgesamtheit (GG), Netto-Stichprobe nach Erstversand (NSP1), Netto-Stichprobe nach Erstversand und 1. Mahnung (NSP2) und gesamter Netto-Stichprobe (NSP3) (in Prozent).

Auch in Bezug auf den höchsten Schulabschluss fällt auf, dass die Befragtenstichprobe gegenüber der Population verzerrt ist: Personen ohne Abschluss bzw. mit Haupt- oder Volksschulabschluss sind unterrepräsentiert, während v. a. Personen mit höheren Bildungsabschlüssen (Fachhochschulreife bzw. Abitur) deutlich überproportional vertreten sind (vgl. Abbildung 2). Wie bereits weiter oben erwähnt, ist dies in Bezug auf den postalischen Befragungsmodus ein häufiger Befund.

Abbildung 2 Rücklaufquoten nach Schulbildung



Schulbildung in Grundgesamtheit (GG), Netto-Stichprobe nach Erstversand (NSP1), Netto-Stichprobe nach Erstversand und 1. Mahnung (NSP2) und gesamter Netto-Stichprobe (NSP3) (in Prozent).

Die Schulbildung unterscheidet sich ferner signifikant zwischen den Erhebungswellen²⁷ (ähnliche Ergebnisse vgl. Reuband 2001: 89f.). Der Anteil derjenigen Personen, die als höchsten Schulabschluss einen Haupt- bzw. Volksschulabschluss oder den Abschluss der Polytechnischen Oberschule/10. Klasse (DDR) aufweisen, steigt mit zunehmenden Mahnaktionen an, während sich die Anteile der Personengruppen mit anderen Abschlüssen verringern.

Die Haushaltsgröße hingegen wird durch die gesamte Befragungstichprobe recht gut abgebildet. Es ergeben sich lediglich kleinere Abweichungen von der Verteilung in der Grundgesamtheit.²⁸

Zusammenfassend kann Folgendes festgehalten werden: ein Vergleich der Einkommens-Mittelwerte in der Grundgesamtheit und allen Netto-Stichproben zeigt, dass der Mittelwert von NSP3 demjenigen der GG noch am stärksten angenähert ist²⁹. Hinsichtlich des höchsten Schulabschlusses sowie der Haushaltsgröße bildet ebenfalls die NSP3 die Verteilung in der GG am besten ab. Die Unterschiede zwischen den einzelnen Netto-Stichproben (v. a. zwischen NSP2 und NSP3) sind allerdings stets marginal. Mit zunehmenden Mahnaktionen ergibt sich demnach keine wesentliche Verbesserung der Stichprobenzusammensetzung in Bezug auf die drei untersuchten Merkmale der Befragten.

Abschließend wurde für drei weitere sozio-demografische Merkmale der Befragten – Berufsbildung, Erwerbstätigkeit³⁰ und subjektiver Sozialstatus³¹ – unabhängig von deren Verteilung in der Grundgesamtheit geprüft, ob diese sich in Abhängigkeit der Erhebungswelle verändern. Während die Erwerbstätigkeit in *keinem* signifikanten Zusammenhang zum Erhebungszeitpunkt steht, unterscheiden sich die

27 Chi-Quadrat=20,838; df=10; Cramers V=0,074; p(2-seitig)=0,022.

28 Während Zwei-Personen-Haushalte überproportional häufig erfasst wurden, sind Ein- und Mehrpersonenhaushalte (ab drei Haushaltsmitglieder) geringfügig unterrepräsentiert.

29 Zur besseren Übersicht wurden in Abb. 1 einige Kategorien zusammengefasst dargestellt. Das Personen-Nettoeinkommen ist ursprünglich wie folgt kodiert: (1) 0-499 €; (2) 500-799 €; (3) 800-999 €; (4) 1.000-1.499 €; (5) 1.500-1.999 €; (6) 2.000-2.999 €; (7) 3.000-über 7.000 €. Es ergeben sich entsprechend folgende Mittelwerte: Mean GG=3,55; Mean NSP1=3,77; Mean NSP2=3,65; Mean NSP3=3,64.

30 „Sind Sie zurzeit erwerbstätig? Unter Erwerbstätigkeit wird jede bezahlte bzw. mit einem Einkommen verbundene Tätigkeit verstanden, egal welchen zeitlichen Umfang sie hat.“ – voll erwerbstätig (min. 35 h/Woche); teilzeitbeschäftigt (15-34 h/Woche); geringfügig erwerbstätig (bis max. 14 h/Woche); gelegentlich oder unregelmäßig beschäftigt; nicht erwerbstätig (Arbeitslose, Vorruheständige/Rentner/Pensionäre ohne Nebenverdienst).

31 „In unserer Gesellschaft gibt es Bevölkerungsgruppen, die eher oben stehen und solche, die eher unten stehen. Sie sehen hier eine Skala, die von 1 bis 10 reicht, wobei 1 das untere Ende und 10 das obere Ende darstellt. Wenn Sie an sich selbst denken, wo auf dieser Skala würden Sie sich einordnen?“ – 1 (ganz unten)...10 (ganz oben). (Vgl. V450/ALLBUS 1990; V724/ALLBUS 2000. Nähere Informationen zum ALLBUS: <http://www.gesis.org/dienstleistungen/daten/umfragedaten/allbus/> (07.06.2010).

Befragten jeder Erhebungswelle hinsichtlich ihrer Berufsbildung³² und ihrer selbst eingeschätzten Stellung innerhalb der Gesellschaft³³ voneinander. Mit zunehmender Kontaktierung werden vorrangig Personen in die Stichprobe rekrutiert, die eine Berufsbildung auf weniger hohem Niveau absolviert haben und sich stärker im unteren Bereich der Gesellschaft einordnen. In Übereinstimmung mit Blasius/Reuband (1996), Reuband (1999b, 2001) sowie Petermann (2005) konnte also auch hier festgestellt werden, dass durch Mahnaktionen besonders Personen unterer sozialer Schichten bzw. Personen mit geringerer Bildung in die Befragung einbezogen werden. Geht man davon aus, dass postalische Befragungen üblicherweise Verzerrungen in Richtung höher Gebildeter aufweisen (vgl. Schnell/Hill/Esser 1995: 334; Reuband 2001), bedeutet die Rekrutierung dieser Gruppierungen durch Mahnaktionen eine Verbesserung der Stichprobenqualität.

5.3 Auswirkungen auf die Teilnahmebereitschaft von Menschen mit verschiedenen Einstellungen, Eigenschaften, Erfahrungen und Verhaltensweisen

Die im vorangegangenen Abschnitt überwiegend durchgeführten „Repräsentanzabgleiche“ mit offiziellen Daten (Nettovalidierung) können selbstverständlich lediglich als Indiz, nicht aber als Beweis für eine gelungene Abbildung der Grundgesamtheit durch die Zufallsstichprobe gelten (vgl. auch Schnell/Hill/Esser 1995: 287). Zahlreiche weitere Merkmale, die in keinem oder nur sehr schwachem Zusammenhang mit den geprüften Sozio-Demografika stehen, sind für die Charakterisierung der Grundgesamtheit bzw. der Stichprobe von Bedeutung. Deshalb sollen in einem nächsten Schritt die Befragten der einzelnen Erhebungswellen hinsichtlich verschiedener weiterer Merkmale miteinander verglichen werden. Da diese Variablen nicht extern validiert werden können, ist zwar aus einem Vergleich der Befragungsteilnehmer „[...] nicht abzulesen, ob bestehende Verzerrungen ausgeglichen, abgeschwächt, verstärkt oder umgekehrt werden“ (Petermann 2005: 70). Es kann aber festgestellt werden, ob sich die Verteilungen der einzelnen Merkmale in Abhängigkeit des Erhebungszeitpunktes verändern. Dies gibt Hinweise darauf, inwiefern die Ausschöpfungsquote bzw. das Durchführen von Mahnaktionen Konsequenzen für die Verteilung der Merkmale bzw. die Generalisierung der Ergebnisse

32 Die Anteile derer, die eine Fachhochschule (Chi-sq=12,405; df=4; Cramers V=0,056; p(2-seitig)=0,015) oder Hochschule (Chi-sq=5,922; df=2; Cramers V=0,054; p(2-seitig)=0,052) absolviert haben, sinken mit zunehmenden Mahnaktionen.

33 Zusammenhang zwischen subjektivem Sozialstatus und Antwortzeitpunkt: F=4,642; p(2-seitig)=0,010, df=2. Die Korrelation ist schwach negativ (Spearman's Rho=-0,078, p≤0,01).

hat. Ein weiterer Vorteil dieser Vorgehensweise besteht ferner darin, dass mehr als nur die wenigen sozio-demografischen Variablen wie Alter, Geschlecht, Bildung etc. zur Analyse potenzieller Stichprobenveränderungen herangezogen werden können.

Zunächst werden Zusammenhänge zwischen verschiedenen Merkmalsausprägungen und den Nachfassaktionen auf bivariater Ebene betrachtet. Die entsprechenden Analyseresultate sollten aufgrund fehlender Drittvariablenkontrolle allerdings unter Vorbehalt und mit Zurückhaltung interpretiert werden.

Tabelle 2 gibt Hinweise darauf, inwiefern sich aus der Durchführung von Mahnaktionen hinsichtlich delinquenzbezogener Angaben Veränderungen in der Stichprobe ergeben.

Tabelle 2 Delinquenzbezogene Antworten nach Erhebungswellen

	Antworten nach Erstversand	Antworten nach 1. Mahnung	Antworten nach 2. Mahnung	Gesamt	Cramers V/ Spearmans Rho
Delinquentes Verhalten					
1) Versatilität, Lebensspanne (5-14 Deliktarten)	16,2	13,4	12,8	14,9	-0,12**
2) Prävalenz Delinquenz, Lebensspanne (2)	72,1	60,1	59,0	66,8	0,13***
3) Prävalenz Delinquenz, seit 50. Geburtstag (2)	51,1	43,3	39,6	47,3	0,09***
4) Opfer v. Diebst./Raub/Betrug, letzte 5 Jahre (2)	7,8	9,6	10,6	8,7	0,04
5) Kriminalitätsfurcht, Standarditem (3 & 4)	11,8	12,3	10,3	11,7	0,03
6) Allgemeine Gesetzestreue (arithm. Mittel)	8,1	8,1	8,2	8,1	0,02

Anteile in Prozent bzw. arithmetisches Mittel; grau: $p \leq 0,01$ in Klammern: dargestellte Antwortkategorien; ** $p \leq 0,01$; *** $p \leq 0,001$; Cramers V bzw. Spearmans Rho beziehen sich stets auf alle Antwortkategorien der Variablen; zu 6) $F=0,280$; $p(2\text{-seit.})=0,76$; $df=2$

Fragewortlaut/Variablen: 1) Es wurde für 14 Deliktarten abgefragt, ob diese im Leben jemals begangen wurden und ein Summenindex zur Anzahl begangener Deliktarten über die Lebensspanne gebildet. 2) Es wurde für 14 Deliktarten abgefragt, ob diese im Leben jemals begangen wurden und über alle Delikte die Lebenszeitprävalenz berechnet: 1) keines der 14 Delikte jemals begangen, lebenslang; 2: mindestens eines der 14 Delikte min. einmal verübt, lebenslang. 3) Es wurde für 14 Deliktarten abgefragt, ob diese seit dem 50. Geburtstag begangen wurden und über alle Delikte die Prävalenz berechnet: 1) keines der 14 Delikte begangen seit dem 50. Geburtstag; 2: mindestens eines der 14 Delikte min. einmal verübt, seit dem 50. Geburtstag. 4) Sind Ihnen die folgenden Ereignisse in den letzten 5 Jahren, also seit Beginn des Jahres 2004, passiert? Ich wurde bestohlen, beraubt oder anderweitig um mein Eigentum betrogen. 1: nein, 2: ja. 5) Wie sicher fühlen Sie sich oder würden Sie sich fühlen, wenn Sie hier in Ihrem Wohngebiet nachts draußen alleine sind? 1: Sehr sicher, 2: Ziemlich sicher, 3: Ziemlich unsicher, 4: sehr unsicher. 6) Bitte geben Sie an, wie sehr Sie persönlich den folgenden Aussagen zustimmen! An Gesetze muss man sich immer halten, auch wenn die eigenen Interessen dabei zu kurz kommen. 1: Stimme überhaupt nicht zu ...10: Stimme voll und ganz zu.

Zwar unterscheiden sich die Befragten in Bezug auf die weniger sensitiven Informationen zu Viktimisierung, Kriminalitätsfurcht und allgemeiner Gesetzestreue nicht signifikant voneinander.³⁴ Angaben zum Schwerpunktthema der Befragung – eigenes kriminelles Verhalten – werden allerdings von den Mahnaktionen beein-

flusst. Es wird offensichtlich, dass sowohl die Versatilität (Deliktbreite) als auch das delinquente Verhalten über die Lebensspanne bzw. seit dem 50. Geburtstag mit zunehmender Kontaktierung abnehmen.³⁵ Unter der Annahme, dass das Auftreten sozial erwünschten Antwortverhaltens vom Antwortzeitpunkt unabhängig ist,³⁶ können die schwächer ausgeprägten Delinquenzangaben als geringerer Bezug der Respondenten späterer Kontaktierungen zum Thema der Befragung interpretiert werden. Damit würde sich hier erneut der Befund früherer Studien bestätigen, wonach thematisch weniger stark interessierte Personen erst durch Nachfassaktionen zur Teilnahme bewegt werden (Reuband 1999b, 2001; Petermann 2005). Die Konsequenzen für das Sample sind eindeutig: hätte man die Befragten lediglich einmal angeschrieben, würde man das kriminelle Verhalten der Population um einige Prozentpunkte³⁷ überschätzen. Zudem liefern die Ergebnisse ein Indiz dafür, dass die Befolgung der Reziprozitätsnorm nicht zwangsläufig mit der Befolgung anderer sozialer (Konformitäts-)Normen zusammenhängt. Anders ausgedrückt: die Vermutung, dass „kriminellere“ Personen in postalischen Befragungen weniger kooperationsbereit sind, also z. B. erst nach (mehrmaliger) Erinnerung antworten, erscheint empirisch unbegründet.³⁸

Tabelle 3 zeigt für weitere „weiche“ Merkmale der Befragten, ob diese von den Erhebungswellen beeinflusst werden. Es ist erkennbar, dass sich die Befragten verschiedener Erhebungswellen hinsichtlich ihres Bedürfnisses nach einfachen Aufgaben (simple tasks), ihrer hedonistischen Orientierung, anomischer Gefühle sowie in Bezug auf ihre Hilfsbedürftigkeit, ihre politische Interessiertheit und Aktivität bzw. Mobilität signifikant voneinander unterscheiden. Die Zusammenhänge sind jedoch stets sehr schwach und teilweise nicht linear. So sind etwa weniger hedonistisch orientierte Befragte sowie der höchste Anteil anomischer Personen in der Stichprobe der 1. Mahnung vertreten. *Keine* Auswirkungen haben die Mahnaktionen auf die Befragtenmerkmale Sorgenanzahl, Lebenszufriedenheit und eigener Lebensstandard³⁹ (vgl. ähnlich Reuband 1999b).

35 Vom Erhebungsdesign unbeeinflusst sind die Anteile derjenigen Personen, die innerhalb der letzten 12 Monate mindestens eines der 14 Delikte verübt haben sowie die Häufigkeiten delinquenten Handelns seit dem 50. Geburtstag bzw. in den letzten 12 Monaten.

36 Die Möglichkeit, dass mit zunehmenden Mahnaktionen der Anteil sozial erwünscht antwortender Personen steigt, kann hier nicht geprüft und somit auch nicht ausgeschlossen werden.

37 Je nach Kriminalitätsmaß würde hier eine Überschätzung um 1,3 bis 5,5 % erfolgen. Diese Werte ergeben sich aus den Differenzen der jeweiligen Anteile „nach Erstversand“ und „gesamt“.

38 Vgl. auch die Ergebnisse der multivariaten Analyse in Tabelle 4.

39 Hinsichtlich drei weiterer reflexiver Wertorientierungen (normorientierte Leistungsethik, religiöse- und subkulturell-materialistische Orientierung), die laut Hermann (2003) in Zusammenhang mit kriminellem Verhalten stehen, sowie hinsichtlich ihres Ausmaßes an Selbstkontrolle, welches nach Gottfredson/Hirschi (1990) kriminalitätsrelevant ist, unterscheiden sich die Befragtenstichproben der einzelnen Erhebungswellen ebenfalls *nicht* signifikant voneinander.

Tabelle 3 Eigenschaften, Wertorientierungen, Einstellungen und Verhaltensweisen nach Erhebungswellen

	Antworten nach Erstversand	Antworten nach 1. Mahnung	Antworten nach 2. Mahnung	Gesamt	Cramers V/ Spearmans Rho
1) Selbstkontrolle: Simple-tasks (arithm. Mittel)	2,6	2,6	2,5	2,6	-0,06**
2) hedonistische Orientierung (arithm. Mittel)	3,6	3,4	3,6	3,6	-0,05*
3) wenige Sorgen (0 & 1)	72,4	74,1	72,5	72,8	0,00
4) starke Anomia (1 & 2)	69,1	76,6	69,8	71,4	0,06*
5) hohe Lebenszufriedenheit (1 & 2)	85,0	81,6	80,8	83,4	0,03
6) guter Lebensstandard (1&2)	87,3	85,7	84,4	86,4	0,04
7) Hilfsbedürftigkeit vorhanden (1)	20,9	26,5	30,7	23,8	0,09***
8) politisches Interesse vorhanden (1-4)	42,8	37,2	36,0	40,2	0,06**
9) hohe Mobilität bzw. Aktivität (1 & 2)	75,0	69,2	65,7	72,0	0,07**

Anteile in Prozent bzw. arithmetisches Mittel; grau: $p \leq 0,05$ in Klammern: dargestellte Antwortkategorien; * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$; Cramers V bzw. Spearmans Rho beziehen sich stets auf alle Antwortkategorien der Variablen; zu 1) $F=3,241$; $p(2\text{-seit.})=0,04$; $df=2$; zu 2) $F=6,221$; $p(2\text{-seit.})=0,002$; $df=2$.

Fragewortlaut/Variablen: 1) Es wurde die Subskala „Simple tasks“ der Self-Control-Skala nach Grasmick et al. (1993) in der dt. Übersetzung von Seipel/Eifler (2004) anhand von vier Items erhoben. Werte der Antwortkategorien liegen zwischen 1 und 4. Nach gleichsinniger Polung der Items bedeuten niedrigere Werte ein höheres Bedürfnis nach einfachen und weniger komplexen Aufgaben. 2) Es wurde die hedonistische Orientierung als individuell reflexive Wertorientierung (genauer: als Wertdimension innerhalb einer Wertorientierung) nach Klages/Gensicke (1993) und Herbert (1993) (zitiert nach Hermann 2003: 58) anhand von vier Items erhoben und zu einem Index zusammengefasst. Werte der Antwortkategorien liegen zwischen 1 und 6. Niedrigere Werte bedeuten eine schwächere hedonistische Orientierung. 3) Es wurde für acht verschiedene Dinge, die Menschen Sorgen bereiten können abgefragt, wie oft die Befragten daran denken. 1: Nie, 2: Selten, 3: Manchmal, 4: Oft, 5: Sehr oft. Diese Variable wurde so rekodiert, dass die Anzahl der Dinge, an welche Personen oft bzw. sehr oft denken, kategorisiert erfasst wurde: 0: keine Sorge; 1: 1 bis 2 Sorgen, 3: 3 bis 4 Sorgen, 4: 5 bis 8 Sorgen. 4) Stimmen Sie diesen Aussagen (eher) zu oder lehnen Sie diese (eher) ab? Heutzutage ändert sich alles so schnell, dass man oft nicht weiß, woran man sich halten soll. 1: Stimme voll zu, 2: Stimme eher zu, 3: Lehne eher ab, 4: Lehne stark ab. 5) Wie beurteilen Sie Ihr Leben insgesamt? Schätzen Sie ein, inwieweit die folgenden Aussagen auf Sie zutreffen! Die wichtigsten Dinge, die ich im Leben will, habe ich weitestgehend erreicht. 1: Trifft genau zu, 2: Trifft eher zu, 3: Trifft eher nicht zu, 4: Trifft gar nicht zu. 6) Wie bewerten Sie Ihren derzeitigen Lebensstandard? 1: Sehr gut, 2: eher gut, 3: eher schlecht, 4: sehr schlecht. 7) Bräuchten Sie mehr Hilfe bei Arbeiten im Haushalt? 1: Ja, 2: Nein. 8) Wenn Sie einmal an die letzten 12 Monate denken, wie oft gehen Sie da in der Regel den folgenden Tätigkeiten nach? Wie oft besuchen Sie politische Veranstaltungen, z. B. von Parteien, Gewerkschaften oder Bürgerinitiativen? 1: Mehrmals in der Woche, 2: Einmal in der Woche, 3: 1-3mal im Monat, 4: Seltener; 5: Nie. 9) Wie oft sind Sie in Ihrem täglichen Leben, egal aus welchen Gründen, auch außerhalb Ihres unmittelbaren Wohngebietes unterwegs? 1: Täglich, 2: 3-6mal pro Woche, 3: 1-2mal pro Woche, 4: 1-3mal im Monat, 5: seltener.

Da die Effekte der Mahnaktionen auf Einstellungen, Wertorientierungen und ähnliche Variablen schwach oder nicht signifikant sind, unterscheiden sich die Befragten des Erstversands nur minimal von denen der gesamten Befragungsstichprobe.⁴⁰ Insgesamt kann daher festgehalten werden, dass die psychosozialen Befindlichkeiten der Befragten durch mehrmalige Kontaktierungen annähernd reproduziert werden (vgl. auch Reuband 1999b: 96).

Allerdings gibt es Zusammenhänge zwischen Mobilität/Aktivität bzw. Anomia und dem Bildungsniveau: je geringer die Schulbildung ist, desto seltener sind die Personen unterwegs und desto häufiger fühlen sie sich anomisch.⁴¹ Ferner ist die Schulbildung mit dem subjektiven Sozialstatus korreliert ($r=0,32$). Um nun genauer zu erkennen, inwiefern die „weichen“ signifikant unterschiedlichen Merkmale untereinander zusammenhängen bzw. von sozio-demografischen Variablen beeinflusst werden und um zu prüfen, inwiefern sich die einzelnen Befragtenstichproben voneinander unterscheiden, werden nun Berechnungen auf multivariater Ebene durchgeführt. Tabelle 4 zeigt die Ergebnisse von drei logistischen Regressionsmodellen der Merkmale von Befragten verschiedener Erhebungszeitpunkte.⁴²

Auffällig ist zunächst, dass einige Zusammenhänge, die in den bivariaten Analysen signifikant sind, in der multivariaten Betrachtung nicht mehr nachgewiesen werden können: dies trifft auf die Schulbildung,⁴³ sowie die Versatilität und Prävalenz kriminellen Handelns seit dem 50. Geburtstag zu. Ein Alterseffekt, der sich auf bivariater Ebene *nicht* zeigt, ist hingegen nun zu beobachten: entsprechend den Befunden bisheriger Untersuchungen und gemäß der Tendenz, die zumindest auch in vorliegender Studie bereits auftritt (vgl. Abschnitt 5.2), ist die Wahrscheinlichkeit späterer Teilnahme für jüngere Personen höher als für ältere.

In Übereinstimmung mit den bivariaten Analysen zeigen sich in jeweils mindestens einem Modell, welchen Einfluss die Wohnortgröße, der subjektive Sozialstatus, das politische Interesse, die Aktivität sowie die frühere Kriminalität eines Befragten auf die Wahrscheinlichkeit ausüben, dass er erst nach ein- bzw. mehrmaliger Mahnung kooperiert.

Des Weiteren geht aus Tabelle 4 hervor, dass sämtliche Modelle, gemessen anhand des Pseudo- R^2 , eine sehr geringe Erklärungskraft aufweisen. Entsprechend geringfügig sind die systematischen Stichprobenverzerrungen, die sich aufgrund der geprüften Variablen ergeben. Mit anderen Worten: die Teilnehmer verschiedener Erhebungszeitpunkte unterscheiden sich hinsichtlich dieser Merkmale nicht wesentlich voneinander. Modell 2 besitzt dabei die höchste Erklärungskraft, d. h. die Teilnehmer des Erstversands und der 2. Mahnaktion unterscheiden sich noch

41 Spearman's Rho (Mobilität bzw. Aktivität/Bildung)=-0,12** und Spearman's Rho (Anomia/Bildung)=0,33***. Letztere Korrelation stellt zugleich die stärkste aller möglichen Zusammenhängebeziehungen aller in die bivariaten Betrachtungen eingeschlossenen Variablen dar.

42 Um die Modelle nicht zu überladen, wurden lediglich einige bedeutsame Variablen in die Berechnung der logistischen Regressionen einbezogen. Nicht aufgenommene Variablen vorangegangener Analysen sind: Haushaltsgröße, Berufstätigkeit, Anomia, hedonistische Orientierung und Simple Tasks.

43 Scheinbar unterscheiden sich die Befragten verschiedener Mahnaktionen stärker bezüglich der subjektiven Einschätzung ihres Sozialstatus als im Hinblick auf ihr formales Bildungsniveau voneinander.

am stärksten – nämlich in Bezug auf sechs Merkmalsverteilungen – voneinander. Allerdings unterscheiden sich auch die Befragten des Erstversands von der gesamten Befragungsstichprobe in fünf Merkmalen voneinander (Modell 3): die Befragten der gesamten Stichprobe nach Durchführung der Mahnaktionen sind im Vergleich zu Personen, die bereits auf das erste Anschreiben reagieren, jünger, wohnen verstärkt im ländlichen Raum, haben geringeres politisches Interesse (bzw. nehmen zu einem größeren Anteil „nie“ an politischen Veranstaltungen teil) und sind im Verlauf ihres Lebens deutlich weniger kriminell. Am ähnlichsten sind sich die Teilnehmer des Erstversands und der 1. Mahnung (Modell 1).

Tabelle 4 Logistische Regressionen der Respondentenmerkmale

Variablen	Teilnahme nach Erstversand vs. nach 1. Mahnung	Teilnahme nach Erstversand vs. nach 2. Mahnung	Teilnahme nach Erstversand vs. nach Erstversand, 1. oder 2. Mahnung
Geschlecht (Ref.: w)	-0,036 (0,123) [0,965]	0,043 (0,158) [1,044]	-0,012 (0,107) [0,988]
Alter (in Jahren)	-0,137* (0,066) [0,872]	-0,192* (0,084) [0,825]	-0,160** (0,058) [0,852]
Schulbildung (Ref.: Fachabitur/Abitur)			
kein SA/Haupt-/Volksschule	0,250 (0,156) [1,284]	0,057 (0,203) [1,059]	0,190 (0,137) [1,209]
Realschule/ 10. Klasse POS (DDR)	0,022 (0,182) [1,022]	-0,068 (0,235) [0,934]	0,009 (,158) [1,009]
Wohnort (Ref.: Dorf)			
Groß-/Mittelstadt	-0,338* (0,157) [0,713]	-0,605* (0,194) [0,546]	-0,423** (0,136) [0,655]
Kleinstadt	-0,230 (0,152) [0,794]	-0,422* (0,187) [0,656]	-0,290* (0,133) [0,748]
subjektiver Sozialstatus	-0,060 (0,062) [,942]	-0,130! (0,078) [0,878]	-0,083 (0,055) [0,920]
politisches Interesse (Ref.: nie)			
oft/manchmal	-0,465 (0,353) [0,628]	-0,393 (0,430) [0,675]	-0,443 (0,296) [0,643]
selten	-0,112 (0,124) [0,894]	-0,317! (0,165) [0,729]	-0,183! (0,109) [0,833]
Aktivität/Mobilität	0,021 (0,065) [1,021]	0,174* (0,078) [1,190]	0,077 (0,056) [1,080]
Versatilität, Lebensspanne	-0,007 (0,079) [0,993]	-0,135 (0,112) [0,874]	-0,047 (0,070) [0,954]
Prävalenz Delinquenz, Lebensspanne (Ref.: min. ein Delikt verübt, Lebensspanne)	0,597*** (0,195) [1,816]	0,391 (0,242) [1,478]	0,522** (0,168) [1,685]
Prävalenz Delinquenz, seit 50. Geburtstag (Ref.: min. ein Delikt verübt, seit 50. Geburtstag)	-0,187 (,164) [,829]	0,016 (,207) [1,016]	-0,117 (0,141) [0,890]
Konstante	-0,741*** (0,164) [,829]	-1,120*** (0,259) [,326]	-0,228 (0,179) [0,796]
Pseudo-R ² Nagelkerke	0,04	0,07	0,05
Chi-Quadrat (df)	38,20*** (13)	49,83*** (13)	62,01*** (13)
n	1411	1205	1643

grau: $p \leq 0,01$

metrische Variablen z-standardisiert; ! $p \leq 0,1$; * $p \leq 0,05$; ** $p \leq 0,01$; *** $p \leq 0,001$

Angegeben sind die unstandardisierten Koeffizienten; in runden Klammern: Standardfehler; in eckigen Klammern: stand. Effektkoeffizienten.

6 Zusammenfassung und Schlussfolgerungen

Der vorliegende Beitrag untersuchte auf der Grundlage einer regionalen postalischen Befragung älterer Menschen die Auswirkungen von Mahnaktionen auf das Antwortverhalten der Respondenten und die soziale Zusammensetzung der Stichprobe.

Die Untersuchung konnte erneut zeigen, dass die postalische Befragung – sofern bei der Konzeption und Durchführung die grundlegenden Empfehlungen der Total/Tailored Design Method (Dillman 1978, 2000; Dillman/Smyth/Christian 2009) beachtet werden – hinsichtlich realisierbarer Teilnahmequoten und der Datenqualität ähnlich erfolgreich wie andere Erhebungsmodi sein kann. Entgegen teilweise bestehender Skepsis stellt sie somit durchaus eine funktionierende Datenerhebungsmethode dar. Ferner scheint der postalische Modus auch für die Befragung älterer Menschen gut geeignet zu sein⁴⁴ (vgl. ähnlich Brune/Werle/Hippler 1991; Reuband 2006, 2008b, 2008a).

Was die Auswirkungen von Mahnaktionen betrifft, konnte Folgendes festgestellt werden:

1. Mahnaktionen erhöhen die Ausschöpfungsquote einer Befragung beträchtlich. Die von Dillman (1978: 185f.) und Hippler (1988: 245) proklamierte „Faustregel“, dass jede Nachfassaktion in etwa die gleiche Ausschöpfungsrate erzielt, konnte allerdings nicht bestätigt werden (ähnliche Befunde vgl. Kunz 2006: 111f.). Die einzelnen Kontaktierungen erbrachten deutlich unterschiedliche, mit dem Zeitpunkt der Mahnung sinkende Rücklaufraten. Drei Erhebungszeitpunkte erscheinen optimal, eine vierte Kontaktierung würde kaum zur Verbesserung der Stichprobenqualität und zur Steigerung des Rücklaufs beitragen.
2. Entsprechend den Ergebnissen bisheriger Studien zeigte sich auch hier, dass das Ausmaß an Item-Nonresponse (INR) mit den Nachfassaktionen ansteigt. Allerdings wurde kein kontinuierlicher Anstieg der Ausfälle einzelner Antworten, sondern ein kurvilinearere Zusammenhang zwischen Mahnung und INR beobachtet: die nach der 1. Mahnung retournierten Bögen weisen mit 7,8 %

44 Je nach Befragungsthema und Umfang der interessierenden Fragen dürfte die postalische Befragungsvariante aufgrund ihrer Spezifika für ältere Menschen sogar besser als andere Erhebungsmodi geeignet sein. Im Rahmen des Pretest der Südbadener Befragung gaben einige Personen (v. a. die älteren) an, dass sie für die Beantwortung des langen Bogens insgesamt mehrere Stunden benötigt hätten. In solchen Fällen „verteilen“ die Befragten ihre Auskunftserteilung auf verschiedene Tagesabschnitte bzw. sogar auf mehrere Tage. Mittels persönlicher oder telefonischer Befragung hätte die hier angestrebte Datenerhebung wahrscheinlich eine massive Überforderung der Befragten dargestellt und wäre entsprechend schwerer zu realisieren gewesen. Zudem enthielt die Befragung zahlreiche Erinnerungsfragen sowie sensitive Fragen zur eigenen Delinquenz – auch hier hätten die beiden letztgenannten Erhebungsmodi sicher zu einer Einschränkung der Antwortvalidität geführt.

den höchsten INR auf. Der Anteil derjenigen Befragten, die den Fragebogen gemeinsam mit einer anderen Person ausfüllen, nimmt mit zunehmender Anzahl der Kontaktierungen linear zu. Der Anstieg der Antwortausfälle und der gemeinschaftlichen Beantwortung ist jedoch jeweils nicht sehr ausgeprägt. Einerseits deuten diese Befunde darauf hin, dass durch Mahnaktionen Personen, die eher unmotiviert sind und solche, deren Fähigkeit zur selbständigen Teilnahme aus verschiedenen Gründen eingeschränkt ist, verstärkt rekrutiert werden. Dies dürfte zu einer verbesserten Stichprobenrepräsentativität führen. Zum anderen wird klar, dass Mahnaktionen eine geringfügige Verminderung der Datenqualität der Befragungsstichprobe bedeuten.

3. Die soziale Zusammensetzung der Stichprobe verändert sich geringfügig mit jeder Nachfassaktion. Vor allem das Alter, der Wohnort, die Schul- und Berufsbildung sowie der subjektive Sozialstatus der Befragten variieren in Abhängigkeit des Teilnahmezeitpunktes. Durch Mahnaktionen werden überwiegend Gruppierungen in die Befragungsstichprobe rekrutiert, die nach einmaligem Anschreiben unterrepräsentiert sind – nämlich Personen, die jünger sind, im ländlichen Raum wohnen, einen schwächeren Sozialstatus aufweisen sowie geringeres politisches Interesse haben. Mahnaktionen tragen damit zu einer Verbesserung der sozialen Zusammensetzung der Stichprobe, d. h. zur Steigerung der Güte der Erhebungsdaten bei. Zwar bildet die gesamte Befragungsstichprobe im Vergleich zur Stichprobe nach dem Erstversand bzw. nach dem Erstversand und nach der 1. Mahnung die Grundgesamtheit hinsichtlich sozialer Merkmale stets am besten ab. Die Unterschiede zwischen der gesamten Befragungsstichprobe und derjenigen nach dem Erstversand sind allerdings insgesamt recht gering.
4. Die Befragten der einzelnen Erhebungszeitpunkte unterscheiden auch bezüglich verschiedener weiterer Merkmale, wie etwa Einstellungen, Verhaltensweisen, Persönlichkeitseigenschaften und Lebenserfahrungen voneinander. Allerdings sind diese Unterscheide überwiegend schwach. Eine Ausnahme stellt das kriminelle Handeln der Befragten über die Lebensspanne dar: hier zeigte sich, dass mit zunehmenden Kontaktierungen deutlich weniger delinquente Personen in die Befragung einbezogen werden. Unterstellt man, dass die später antwortenden Personen (etwa aufgrund eines geringeren Bezugs zum Befragungsthema) in der Befragungsstichprobe des Erstversands unterrepräsentiert sind, würde die Berechnung kriminalitätsbezogener Parameter der Population auf der Basis der Befragten des Erstversands entsprechend zu spürbaren Überschätzungen führen. Auch hier wird noch einmal deutlich, inwiefern Mahnaktionen zur Verbesserung der Stichprobenqualität beitragen können.

Zusammenfassend deuten die Befunde darauf hin, dass auch postalische Befragungen mit geringen Ausschöpfungsquoten (zumindest im hier untersuchten Bereich, d. h. in Höhe von ca. 30 %) bzw. ohne Mahnaktionen eine hinreichend gute Stichprobenqualität im Sinne der Abbildhaftigkeit sozio-demografischer Merkmale der

Grundgesamtheit aufweisen können. Allerdings unterscheiden sich die Befragten ohne und nach weitere(n) Kontaktierungen in Bezug auf mehr oder weniger bedeutsame Variablen mehr oder weniger stark voneinander. Im günstigsten Fall sind bedeutsame Variablen nicht oder nur in geringfügigem Ausmaß von Antwortdifferenzen betroffen. Um die für statistische Analysen benötigte Fallzahl zu erreichen, wäre es unter diesen Umständen theoretisch gerechtfertigt, eine hinreichend große Brutto-Stichprobe einmalig anzuschreiben, d. h. auf Mahnaktionen zu verzichten. Im ungünstigeren Fall – und dies trifft auf die hier untersuchte Erhebung zu und wurde auch für andere Studien bereits nachgewiesen – stehen die Antworten der Respondenten zum Schwerpunktthema der Befragung in deutlichem Zusammenhang zum Erhebungszeitpunkt. Hier können Nachfassaktionen zur Abschwächung anfänglicher Verzerrungen beitragen.

Da im Vorfeld einer Befragung kaum bestimmbar ist, ob die angeschriebenen Personen unabhängig ihres Interesses am Erhebungsthema nach einmaliger Kontaktierung zu je etwa gleichen Anteilen antworten werden, *erscheint es empfehlenswert, Mahnaktionen prinzipiell in das Erhebungsdesign postalischer Befragungen einzuplanen*. Die Ergebnisse vorliegender Studie legen dabei nahe, dass bereits *eine* Nachfassaktion deutlich zur Verbesserung der Stichprobenzusammensetzung und damit zur Steigerung der Befragungsgüte beiträgt. *Zwei* Mahnaktionen sind in dieser Hinsicht optimal. *Drei* Mahnaktionen scheinen in Anbetracht von Kosten-Nutzen-Abwägungen wenig sinnvoll zu sein.

Indem eine geringere Anzahl von Personen kontaktiert wird, diese jedoch besser ausgeschöpft werden kann, wird zugleich der Pool potenziell zu befragender Personen in der Bevölkerung geschont. Ein solches Design stellt eine sozial verantwortungsbewusste Vorgehensweise dar und ist – da keine nachteiligen Auswirkungen von Mahnaktionen festgestellt werden konnten – somit auch aus forschungsethischer Sicht zu bevorzugen.

Literatur

- Arzheimer, K. und M. Klein, 1998: Die Wirkung materieller Incentives auf den Rücklauf einer schriftlichen Panelbefragung. *ZA-Information* 43: 6-43. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/za_information/ZA-Info-43.pdf (9.9.2010).
- Atteslander, P., 2008: *Methoden der empirischen Sozialforschung*. Berlin: Schmidt.
- Becker, R., R. Imhof und G. Mehlkop, 2007: Die Wirkung monetärer Anreize auf den Rücklauf bei einer postalischen Befragung und die Antworten auf Fragen zur Delinquenz. *Methoden Daten Analysen* 1 (2): 131-160. www.gesis.org/download.php?url=/fileadmin/upload/forschung/publikationen/zeitschriften/mda/Vol.1_Heft_2/MDA2_07_Becker_Imhof_Mehlkopf.pdf (9.9.2010).

- Biemer, P. P. und L. E. Lyberg, 2003: *Introduction to Survey Quality*. Hoboken, New Jersey: John Wiley & Sons.
- Blasius, J. und K.-H. Reuband, 1996: Postalische Befragungen in der empirischen Sozialforschung. Ausschöpfungsquoten und Antwortqualität. *Planung & Analyse* 96 (1): 35-41.
- Blau, P., 1964: *Exchange and Power in Social Life*. New York: Wiley.
- Brambilla, D. J. und S. M. McKinlay, 1987: A Comparison of Responses to Mailed Questionnaires and Telephone Interviews In a Mixed Mode Health Survey. *American Journal of Epidemiology* 126 (5): 962-971.
- Brune, M., M. Werle und H.-J. Hippler, 1991: Probleme bei der Befragung älterer Menschen. Methodische Erfahrungen aus einer schriftlichen Umfrage zu Tätigkeitsformen im Ruhestand. ZUMA-Nachrichten 28: 73-91. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_28.pdf (9.9.2010).
- Converse, P., 1964: The Nature of Belief-Systems in Mass Publics. S. in: D. A. Apter (Hg.): *Ideology and Discontent*. New York, London: The Free Press.
- De Leeuw, E. D., 1992: *Data Quality in Mail, Telephone and Face-to-Face Surveys*. Amsterdam: TT-Publikaties.
- Diekmann, A., 1998: *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen*. Reinbek bei Hamburg: Rowohlt.
- Diekmann, K., 2007: *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen*. Reinbek bei Hamburg: Rowohlt.
- Dillman, D. A., 1978: *Mail and Telephone Surveys*. New York: Wiley.
- Dillman, D. A., 2000: *Mail and Internet Surveys: The Tailored Design Method*. New York: Wiley.
- Dillman, D. A., J. D. Smyth und L. M. Christian, 2009: *Internet, Mail, and Mixed-Mode Surveys. The Tailored Design Method*. Hoboken; New Jersey: John Wiley & Sons.
- Dommeyer, C. J., 1985: Does Response to an Offer of Mail Survey Results Interact with Questionnaire Interest? *Journal of Market Research Society* 27 (1): 27-38.
- Friedrichs, J., 1990: *Methoden empirischer Sozialforschung*. Opladen: Westdeutscher Verlag.
- Gottfredson, M. R. und T. Hirschi, 1990: *A General Theory of Crime*. Stanford: Stanford University Press.
- Goudy, W., 1978: Interim Response to a Mail Questionnaire: Impacts on Variable Relationships. *Sociological Quarterly* 19 (2): 253-265.
- Grasmick, H., C. Tittle, R. Bursik und B. Arneklev, 1993: Testing The Core Empirical Implications of Gottfredson and Hirschi's General Theory of Crime. *Journal of Research in Crime and Delinquency* 30 (1): 5-29.
- Groves, R., 1989: *Survey Errors and Survey Costs*. New York: John Wiley & Sons.
- Guadagnoli, E. und S. Cunningham, 1989: The Effects of Nonresponse and Late Response on a Survey of Physician Attitudes. *Evaluation & the Health Professions* 12: 318-328.
- Häder, M., 2006: *Empirische Sozialforschung. Eine Einführung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Herbert, W., 1993: *Wandel und Konstanz von Wertstrukturen*. Frankfurt Main u. a.: Peter Lang.
- Hermann, D., 2003: *Werte und Kriminalität. Konzeption einer allgemeinen Kriminalitätstheorie*. Wiesbaden: Westdeutscher Verlag.
- Hippler, H.-J., 1988: Methodische Aspekte schriftlicher Befragungen: Probleme und Forschungsperspektiven. *Planung & Analyse* 88 (6): 244-248.
- Hippler, H.-J. und K. Seidel, 1985: Schriftliche Befragungen bei allgemeinen Bevölkerungstichproben - Untersuchungen zur Dillmanschen „Total Design Method“. ZUMA-Nachrichten 16: 39-56. www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_16.pdf (9.9.2010).
- Homans, G. C., 1961: *Social Behavior. Its Elementary Forms*. New York: Harcourt, Brace & World.
- Kanuk, L. und C. Berenson, 1975: Mail Surveys and Response Rates: A Literature Review. *Journal of Marketing Research* 12 (Nov.): 440-453.
- Klages, H. und T. Gensicke, 1993: *Erläuterung der Speyerer Ziele und Methodik der Wert-erfassung*. Speyer. unveröff. Manuskript.

- Koch, A., 1998: Wenn „mehr“ nicht gleichbedeutend mit „besser“ ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen. ZUMA-Nachrichten 42: 66-93. http://www.gesis.org/download.php?url=/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_42.pdf (9.9.2010).
- Koenig, D., G. Martin und L. Seiler, 1977: Response Rates and Quality of Data: A Reexamination of the Mail Questionnaire. *Canadian Review of Sociology and Anthropology* 14: 432-438.
- Kunz, F., 2006: Täter-Opfer-Ausgleich im Jugendbereich - wie erleben und bewerten ihn die Konfliktbeteiligten? Eine schriftliche Befragung von Tätern und Opfern aus drei sächsischen Schlichtungsstellen. Institut für Soziologie, TU Dresden, unveröff. Diplomarbeit.
- Kunz, F., 2007: Im Osten was Neues: Täter-Opfer-Ausgleich aus Sicht der Beteiligten. Ergebnisse einer Befragung von jugendlichen bzw. heranwachsenden Tätern und ihren Opfern. *Monatsschrift für Kriminologie und Strafrechtsreform* 90 (6): 466-483.
- Lamnek, S. und R. Trepl, 1991: Die Nichtteilnahme an sozialwissenschaftlichen Befragungen. *Planung & Analyse* 6: 205-211.
- Leslie, L. L., 1972: Are High Response Rates Essential to Valid Surveys? *Social Science Research* 1 (3): 323-334.
- Mehlkop, G. und R. Becker, 2007: Zur Wirkung monetärer Anreize auf die Rücklaufquote in postalischen Befragungen zu kriminellen Handlungen. *Methoden Daten Analysen* 1 (1): 5-24. www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/mda/Vol.1_Heft_1/MDA1_Mehlkop_Becker.pdf (9.9.2010).
- Niedermayer, O., 1989: Zur Übertragung der „Total Design Method“ auf schriftliche Befragungen in der Bundesrepublik. S. 332-348, in: J. W. Falter, H. Rattinger und K. G. Troitzsch (Hg.): *Wahlen und politische Einstellungen in der Bundesrepublik Deutschland*. Frankfurt/Main: Lang.
- Petermann, S., 2005: Rücklauf und systematische Verzerrungen bei postalischen Befragungen. Eine Analyse der Bürgerumfrage Halle 2003. ZUMA-Nachrichten 29 (57): 56-78. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_57.pdf (9.9.2010).
- Platek, R. und C.-E. Särndal, 2001: Can a Statistician Deliver? *Journal of Official Statistics* 17 (1): 1-20.
- Prüfer, P. und M. Rexroth, 2000: Zwei – Phasen – Pretesting. ZUMA-Arbeitsbericht 2000/08. www.gesis.org/download.php?url=/fileadmin/upload/forschung/publikationen/gesis_reihen/zuma_arbeitsberichte/00_08.pdf (9.9.2010).
- Reuband, K.-H., 1999a: Anonyme und nicht-anonyme postalische Bevölkerungsbefragungen. *Planung und Analyse* (1): 56-59.
- Reuband, K.-H., 1999b: Postalische Befragung in den neuen Bundesländern. Durchführungsbedingungen, Ausschöpfungsquoten und soziale Zusammensetzung der Befragten in einer Großstadtstudie. *ZA-Informationen* 45: 71-99. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/za_information/ZA-Info-45.pdf (9.9.2010).
- Reuband, K.-H., 2001: Möglichkeiten und Probleme des Einsatzes postalischer Befragungen. *KZfSS* 53 (2): 307-333.
- Reuband, K.-H., 2006: Postalische Befragung alter Menschen. Kooperationsverhalten, Beantwortungsstrategien und Qualität der Antworten. *ZA-Informationen* 59: 100-127. www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/za_information/ZA-Info-59.pdf (9.9.2010).
- Reuband, K.-H., 2008a: Einsam und unglücklich im Alter? Eine empirische Studie zur Lebenssituation alter Menschen im großstädtischen Kontext. S. 354-374, in: A. Groenemeyer und S. Wieseler (Hg.): *Soziologie sozialer Probleme und sozialer Kontrolle* Wiesbaden: VS Verlag Sozialwissenschaften.
- Reuband, K.-H., 2008b. *Mail Surveys Among the Very Old. A Study of Retrospective Accounts of the Nazi Past*. Paper präsentiert auf der 62. Jahrestagung der American Association for Public Opinion Research (AAPOR), New Orleans.

- Schneekloth, U. und I. Leven, 2003: Woran bemißt sich eine „gute“ allgemeine Bevölkerungsumfrage? Analysen zu Ausmaß, Bedeutung und zu den Hintergünden von Nonresponse in zufallsbasierten Stichprobenerhebungen am Beispiel des ALLBUS. ZUMA-Nachrichten 53: 16-57. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_53.pdf (9.9.2010).
- Schnell, R., 1997: Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen. Opladen: Leske + Budrich.
- Schnell, R., P. B. Hill und E. Esser, 1995: Methoden der empirischen Sozialforschung. München, Wien: Oldenbourg.
- Schnell, R., P. B. Hill und E. Esser, 2008: Methoden der empirischen Sozialforschung. München, Wien: Oldenbourg.
- Scott, C., 1961: Research on Mail Surveys. *Journal of the Royal Statistical Society*. 124 (2): 143-205.
- Seipel, C. und S. Eifler, 2004: Gelegenheiten, Rational-Choice und Selbstkontrolle. Zur Erklärung abweichenden Handelns in High-Cost- und Low-Cost-Situationen. S. 288-316, in: D. Oberwittler und S. Karstedt (Hg.): *Soziologie der Kriminalität*. Wiesbaden: VS Verlag Sozialwissenschaften.
- Sobal, J. und K. C. Ferentz, 1989: Comparing Physicians' Responses to the First and Second Mailings of a Questionnaire. *Evaluation & the Health Professions* 12 (3): 329-339.
- Stoop, I., 2005: *The Hunt for the Last Respondent. Nonresponse in Sample Surveys*. The Hague: Social and Cultural Planning Office.
- Toops, T. A., 1926: The Returns for Follow-Up Letters to Questionnaires. *Journal of Applied Psychology* 10: 92-110.
- Yammarino, F. J., S. J. Skinner und T. L. Childers, 1991: Understanding Mail Survey Response Behavior. *Public Opinion Quarterly* 55: 613-639.
- Yu, J. und H. Cooper, 1983: A Quantitative Review of Research Design Effects on Response Rates to Questionnaires. *Journal of Marketing Research* 20: 36-44.

Anschrift der Autorin Franziska Kunz
Max-Planck-Institut für ausländisches und
internationales Strafrecht
Kriminologische Forschungsgruppe
Günterstalstr. 73
79100 Freiburg i. Br.
f.kunz@mpicc.de