

Kommentar zu Anna Schnauber und Gregor Daschmann: "States oder Traits? Was beeinflusst die Teilnahmebereitschaft an telefonischen Interviews" (MDA 2008, 2: 97–123)

Bock, Olaf; Schnapp, Kai-Uwe

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Bock, O., & Schnapp, K.-U. (2009). Kommentar zu Anna Schnauber und Gregor Daschmann: "States oder Traits? Was beeinflusst die Teilnahmebereitschaft an telefonischen Interviews" (MDA 2008, 2: 97–123). *Methoden, Daten, Analysen (mda)*, 3(2), 249-259. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-127208>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kommentar zu Anna Schnauber und Gregor Daschmann:

*„States oder Traits?
Was beeinflusst die
Teilnahmebereitschaft an
telefonischen Interviews“*
(MDA 2008, 2: 97–123)

Comment to Anna Schnauber and Gregor Daschmann:

*„States or Traits?
Factors Influencing the
Willingness to Participate
in Telephone Surveys“*
(MDA 2008, 2: 97–123)

Olaf Bock und Kai-Uwe Schnapp

1 Einleitung

Das zentrale Ergebnis der Non-Response-Studie von Schnauber und Daschmann (im Folgenden S/D) lautet:

„Die Umfrageeinstellung [...] ist ein stabiler Einflussfaktor. Da sich aber nur wenige und schwache Zusammenhänge mit grundlegenden Persönlichkeitseigenschaften und soziodemografischen Merkmalen zeigen, spricht dies zwar dafür, dass es bestimmte Personen gibt, die Befragungen gegenüber grundsätzlich abgeneigt sind, diese sich aber nicht grundlegend von den Teilnehmern einer Befragung unterscheiden. Somit kann davon ausgegangen werden, dass Verweigerungen *nicht* zu systematischen Verzerrungen der Ergebnisse von Umfragen führen.“ (Schnauber/Daschmann 2008: 97; Hervorhebungen: O.B./K.–U.S.)

Diese Schlussfolgerung hat, wird sie in der Praxis ernst genommen, weitreichende Konsequenzen für die Qualitätssicherung bei (telefonischen) Umfragen. Sie erscheint uns allerdings nicht so gut begründet, als dass sie unbefragt Bestand haben sollte. S/D's Konklusion basiert *erstens* vorrangig auf Aussagen zum fehlenden Zusammenhang zwischen „grundlegenden Persönlichkeitsmerkmalen“ auf der einen und sozialstrukturellen Variablen sowie dauerhaften Einstellungen zu Umfragen auf der anderen Seite. Dieser Nichtzusammenhang wird nach unserer Lesart a) nicht nachgewiesen und wäre b) selbst bei einem Nachweis keine hinreichende Absicherung für die oben getroffene Feststellung. Diese Position werden wir begründen.

Zweitens wird die Frage der Nichterreichbarkeit im Aufsatz gar nicht thematisiert. S/D begründen dies im ersten Absatz damit, dass die Gruppe der Verweigerer „besonders interessant“ für die Nonresponseforschung wäre (S/D 2008: 98). Wir werden erläutern, warum wir die Exklusion der Nichterreichbaren für unangemessen halten. *Drittens* erschwert das Fehlen einer Darstellung zum Verhältnis zwischen Panel- und Erstbefragten in der Studie deutlich die Interpretation der Ausschöpfungsquoten. Darauf möchten wir vor allem deshalb hinweisen, weil S/D die hohe Ausschöpfungsquote ihrer Verweigererstudie im Vergleich zu „Studien in der Vergangenheit“ besonders positiv hervorheben (S/D 2008: 111). Zunächst soll aber der für die Diskussion zentrale Begriff der „Nonresponse“ formal definiert werden.

2 Definition von Non-Response

Fixpunkt der von S/D aufgegriffenen Debatte ist der sogenannte *Nonresponse-Bias*. Er setzt sich zusammen aus dem *Nonresponse-Error*, dem Unterschied zwischen Respondenten R und Nichtrespondenten N , sowie der *Nonresponse-Rate*, dem Anteil der Nonresponse an der Gesamtstichprobe (Schnell 2008: 12). Für ein arithmetisches Mittel kann der *Nonresponse-Error* dann z. B. wie folgt formuliert werden:

$$Bias_{\bar{x}} = (\bar{x}_R - \bar{x}_N) * \frac{n_N}{n_N + n_R}$$

Träfe S/D's Studienergebnis zu – Respondenten (R) und Nichtrespondenten (N) unterscheiden sich in Umfragen *nicht* systematisch (S/D 2008: 97, 120) – so wäre stets von einem *Nonresponse-Error* von Null auszugehen. Das würde das alarmierte Eingangsplädoyer der Autoren (vgl.: 98) für eine dringende Erhöhung der Stichprobenausschöpfungen zur Verringerung der *Nonresponse-Rate* überflüssig machen, denn eine Verringerung der *Nonresponse-Rate* verändert die Untersuchungsdaten und -ergebnisse nicht, wenn ein *Nonresponse-Error* gar nicht vorhanden, der Unterschied zwischen Respondenten und Nichtrespondenten also Null ist. Lässig würden auch neuere elaborierte Methoden zur *Korrektur von Nonresponse-Bias* (siehe dazu Schnell/Hill/Esser 2005: 314ff.; Schnell 2008). Wir halten es jedoch weder für plausibel, von einem über unterschiedliche Umfragestudien einheitlichen noch gar von einem *Nonresponse-Fehler* von Null auszugehen. Diese Position begründen wir im Folgenden.

3 Zusammenhang zwischen „grundlegenden Persönlichkeitsmerkmalen“ und anderen Variablen

S/D stellen zunächst auf der Basis ihrer Daten fest, dass „sowohl bestimmte Traits als auch bestimmte States einen Einfluss auf die Teilnahmebereitschaft“ (119) haben. Weiterhin beobachten sie, dass weder Traits noch States „mit grundlegenden Persönlichkeitsmerkmalen zusammenhängen“ (119). Diese „grundlegenden Persönlichkeitsmerkmale“ werden in der Analyse von S/D über die Variablen „Vertrauen in Mitmenschen“, „Extrovertiertheit“ und „Kommunikation/Kompetenz“ (114f., 118f.) operationalisiert. Die Analyse mündet in der zunächst auf ihre Studie (*Exba 2007*) bezogenen Feststellung, dass aufgrund des fehlenden Zusammenhanges zwischen grundlegenden Persönlichkeitsmerkmalen auf der einen, sozialstrukturellen Variablen und dauerhaften Umfrageeinstellungen auf der anderen Seite, keine grundlegenden Repräsentativitätsprobleme zu befürchten sind.

Um die Aussage eines fehlenden Zusammenhanges zwischen diesen Variablen und z. B. den sozialstrukturellen Variablen zu belegen, reicht es u. E. aber nicht, eine logistische Regressionsanalyse durchzuführen, wie sie S/D in diesem Fall präsentieren, da hier lediglich die Effekte der erklärenden Variablen auf die Befragungsteilnahme geprüft werden. Vielmehr ist die Korrelation innerhalb der Gruppe der erklärenden Variablen explizit zu prüfen.¹ Eine solche Analyse wird in dem Beitrag jedoch nicht vorgelegt. Neben diesem formalen Problem ist aber vor allem auf den substanziellen Punkt hinzuweisen, dass das Vorliegen oder Fehlen von Korrelationen zwischen zwei Variablengruppen (hier „grundlegende Persönlichkeitsmerkmale“ und „sozialstrukturelle Variablen“) keine Auskunft darüber liefert, ob und wie diese Variablengruppen mit anderen Variablengruppen korreliert sind, wie es also um die Eignung der Daten für Inferenzschlüsse in nicht betrachteten Variablensegmenten bestellt ist.

Wenn es zutreffen soll, dass Verweigerungen nicht zu systematischen Verzerrungen bei Umfragen führen, muss man unterstellen, dass S/D von einem vollständig zufälligen Ausfallprozess ausgehen, die Ausfälle wären also „missing completely at random“ (MCAR). Diese Annahme liegt nahe, weil S/D unkonditional von einem Fehlen systematischer Verzerrungen sprechen. Sie müssen also davon ausgehen, dass die Ausfälle weder von den beobachteten noch von den unbeobachteten

1 Die Notwendigkeit der Überprüfung von Korrelationen zwischen den erklärenden Variablen ergibt sich natürlich auch aus der Notwendigkeit, das Fehlen von Multikollinearität als einer Anwendungsvoraussetzung der Regressionsanalyse nachzuweisen. Über einen entsprechenden Test wird in dem Beitrag jedoch nicht berichtet.

Fällen abhängen. Die Ausfälle wären dann komplett ignorierbar, ihr Einschluss in die Analyse würde die Analyseergebnisse nicht ändern (vgl. Weins 2006: 207f.). Wie Weins jedoch weiter feststellt, ist der Ausfallmechanismus in der Regel unbekannt und die Annahme ignorierbarer Ausfälle nicht prüfbar (ebd. 208). Den Beweis der Ignorierbarkeit der Ausfälle treten S/D folgerichtig auch nicht an. Vielmehr rekurren sie auf die Unterstellung, dass eine Homomorphie der Verteilung bestimmter (sozialstruktureller) Variablen zwischen den Teildatensätzen der Teilnehmer und der Verweigerer auf eine Homomorphie der Verteilungen in allen analytisch interessierenden Variablen schließen lässt. Anders ausgedrückt ist die Unterstellung von S/D, dass die Repräsentativität der Stichprobe und damit ihre Eignung für Inferenzschlüsse an keiner Stelle beeinträchtigt ist. Diese Form eines korrelationsbezogenen „Analogieschlusses“ ist jedoch nicht zulässig (Schnell 1993: 29; Gabler et al. 1994). Es muss vielmehr davon ausgegangen werden, dass eine strukturelle Übereinstimmung von Datensätzen in einem Variablenbereich nichts Ausreichendes über eine äquivalente Übereinstimmung auf anderen Variablen aussagt. Darüber hinaus zeigt die Nonresponseforschung, wie oben erläutert, dass diese Übereinstimmung in der Regel auch gar nicht gegeben ist. Nimmt man diese Einwände ernst, dann kann also allenfalls von einem Ausfallmechanismus ausgegangen werden, bei dem die Ausfälle „missing at random“ (MAR) sind. Hier dürfen die Ausfälle zwar von den beobachteten, nicht aber von den unbeobachteten Werten abhängen (Weins 2006: 207f.). Auch den Beweis für die Richtigkeit dieser Annahme treten S/D aber nicht an und können ihn auch nicht antreten (vgl. erneut Weins 2006: 208). Ist ein Nachweis also nicht möglich, dass die Ausfälle MCAR oder MAR sind, müsste man im Sinne einer konservativen Strategie der Qualitätssicherung von einem systematischen Ausfallprozess ausgehen („missing not at random“ – MNAR). Ein solcher Ausfallprozess erlaubt jedoch keine inferenzstatistische Nutzung der Daten. Zusammengefasst lautet unser Argument also: S/D unterstellen implizit einen MCAR-Ausfallmechanismus, dessen Existenz sie aber nicht nachweisen (können). Dieser fehlende Nachweis macht ihre allgemeine Schlussfolgerung, dass die Ausfälle in ihrer speziellen Studie und darüber hinaus in Umfragen generell ignorierbar seien, unglaubwürdig. Deskriptive Aussagen für die erhobenen Daten bleiben damit natürlich möglich. Die Zulässigkeit von Inferenzschlüssen auf der Basis der vorliegenden Daten muss jedoch angezweifelt werden.

4 States und Traits

Neben der Diskussion über die Repräsentativität der Daten ist es nach unserer Ansicht geboten, die Frage zu stellen, welche Aussagen mit den Daten von S/D auf der deskriptiven Ebene über die Rolle von States und Traits, also Eigenschaften der Situation und Eigenschaften der Personen aufgrund der Analyseergebnisse von S/D gemacht werden können.

In der deutschen (Schnell 1997, 2008) wie internationalen (Groves/Couper 1998) Forschung zur Umfragebeteiligung wird seit geraumer Zeit betont, dass die Beteiligung vor allem eine Funktion der Situation, weniger aber fester Überzeugungen für oder gegen die Teilnahme an Umfragen sei:

„We believe that few householders have strongly preformed decisions about survey requests. Rather, these decisions are made largely at the time of the request for participation.“

Nach den Ergebnissen einer Sekundäranalyse des Teilnahmeverhaltens in mehreren Bevölkerungsumfragen sind die Einflüsse auf die Teilnahmeentscheidung „highly susceptible to situational factors.“ (Groves/Couper 1998: 321) Im Sinne der Nomenklatur von S/D sind es also eher States (situative Faktoren) denn Traits (situationsübergreifend invariante Faktoren), die Einfluss auf die Teilnahmeentscheidung an Umfragen haben. Auch andere Arbeiten dokumentieren die Nichtangemessenheit der Vorstellung von der Existenz eines fixen Bevölkerungsstratums so genannter „harter“ Verweigerer (von der Heyde 2002: 39; Niemann/Abel 2000: 119f.).

Diese Ergebnisse der empirischen Beteiligungsforschung werden von S/D nur zum Teil bestätigt. So erklären die situationsübergreifend stabilen Umfrageeinstellungen (vgl. die Systematisierung von S/D: 101) im Regressionsmodell für die Teilnahmeentscheidung einen sehr großen Teil der Teilnahmeentscheidung (vgl. die in der Regressionstabelle auf S. 118 berichteten Pseudo-R²-Werte). Gleichzeitig hat nur ein Teil der situativen Faktoren, nämlich die hier als Stress- und Zeitfaktor deklarierten Variablen einen hohen Erklärungsanteil, während Interviewermerkmale sowie Dauer und Thema des Interviews als weitere Bündel situativer Faktoren nur wenig Erklärungskraft beitragen. Diesen Umstand versuchen S/D vor allem in Bezug auf den geringen gemessenen Einfluss der Interviewer argumentativ abzuschwächen. Das mutet aus zwei Gründen eigenartig an. Zum *einen* könnte man jedes Ergebnis einer konkreten Datenanalyse mit der Begründung nicht ernst nehmen, dass man aus anderer Forschung wisse, dass dieser Faktor einflussreich sei. Zum *anderen* hat die Variablengruppe der sozialstrukturellen Indikatoren eine im Vergleich viel höhere Erklärungskraft, wird aber in der Interpretation der Daten eher abschwächend behandelt. Berücksichtigt man das von Kozak (2009) vorge-

brachte Argument, dass man einen Varianzaufklärungsanteil nicht vor absoluten, sondern vor an den Hypothesen orientierten Schwellwerten interpretieren sollte, dann ist dieses Vorgehen von S/D um so weniger plausibel. Vor dem Hintergrund des Forschungsstandes würde man für die stabilen Merkmale eine sehr geringe, für die situativen Merkmale eine sehr hohe Varianzaufklärung erwarten. Die vorgelegten Ergebnisse weisen aber eine entgegengesetzte Struktur auf.

Diese Abweichung der eigenen Ergebnisse vom Forschungsstand wird von den Autoren nach unserer Einschätzung nicht ausreichend thematisiert. Vielmehr mündet die Analyse etwas voreilig einerseits in einer Entwarnung (keine Zusammenhänge zwischen diesen Umfrageeinstellungen und grundlegenden Persönlichkeitsmerkmalen, daher auch kein Zusammenhang des Ausfalls mit jedweden anderen interessierenden Merkmalen) und andererseits in der aktiven Forderung, dass es vor allem darum gehen müsse, das Image der Umfrageforschung zu verbessern, um höhere Ausschöpfungsquoten und damit eine höhere Repräsentativität von Umfragen zu erreichen (119f.). Letztere Forderung erscheint zwar insofern plausibel, als sie eine Kur für das Problem der ausfallbedingten Verzerrungen empfiehlt, die tatsächlich in der Hand derjenigen liegt, die Umfragen durchführen. Vor dem Hintergrund der oben geführten Argumentation überzeugt uns dieser Lösungsansatz allerdings nicht.

5 Nichterreichbarkeit

Ein weiteres, aus unserer Sicht bedeutsames Manko der Arbeit von S/D ist die grundsätzliche Nichtbefassung mit den nicht erreichbaren Mitgliedern der Bruttostichprobe. Nach einem ersten generellen Bezug auf Ausschöpfungsquoten beschränken sich S/D auf die „besonders interessante Gruppe der Verweigerer“ (98). Diese Einschränkung bei der Untersuchung von Unit-Nonresponse erscheint uns als nicht plausibel, denn für einen Teil der Fachdiskussion ist das Problem der Nichterreichbarkeit inzwischen wesentlich gravierender, als das Problem der Verweigerung (vgl. u. a. Atrostic et al. 2001; Steeh et al. 2000; Baur 2006: 164ff.; Niemann/Abel 2000: 120; Dethlefsen 2000: 63; Tuckel/O'Neill 2002). Die schwindende Teilnahme an Telefonumfragen veranlasst einige Autoren sogar, von einer Transitionsphase zu sprechen, „in which the telephone survey is losing its status as the most popular mode of data gathering.“ (Tuckel/O'Neill 2002: 34) Indem „States oder Traits?“ das Problem der Nichterreichbarkeit völlig unbeachtet lässt, bleibt ein Teil der Unit-Nonresponse in dieser Studie vollkommen unaufgeklärt. Das sehen wir als problematisch an, weil der Ausfall nichterreichbarer Personen nicht zufällig (also mindes-

tens MAR), sondern abhängig von spezifischen Attributen dieser Personengruppe ist (Schnell et al. 2008: 311). Frühere Nonresponsestudien haben gezeigt, dass sich leicht und schwer Erreichbare deutlich unterscheiden, so dass der Ausfall von schwer Erreichbaren eine Stichprobe erheblich verzerren kann (Baur 2006: 169). Unklar bleibt bei den von S/D gemachten Angaben zum Stichprobendesign (107, Anm. 5), ob im Rahmen der Hauptstudie überhaupt mehrere Kontaktversuche bei allen ausgewählten Stichprobenelementen unternommen wurden oder nicht vielmehr Nichtkontakte durch „neue“ Stichprobenelemente ersetzt wurden. Im ersten Fall wäre die unzureichende Dokumentation der Ausfälle zu kritisieren, im zweiten läge einerseits ein problematisches, dem Redressment nahes Auswahlverfahren der Studie zugrunde, das Inferenzschlüsse problematisch macht. Mit redressmentähnlichen Ziehungsverfahren werden aber Nichterreichbarkeitsquoten über dem Erwartungswert erzeugt, weil vor allem leicht erreichbare Personen in die Stichprobe gelangen. Schwer erreichbare Personen haben eine deutlich unter dem Erwartungswert von Studien mit mehrfachen Kontaktversuchen liegende Wahrscheinlichkeit, überhaupt in die Stichprobe zu gelangen. Die oben zitierten strukturellen Unterschiede zwischen leicht und schwer erreichbaren Personen führen dann zu einer gegenüber der Grundgesamtheit verzerrten Stichprobe. Da S/D die Leserschaft nicht ausreichend über das Stichprobendesign aufklären², bleibt zunächst ungeklärt, ob neben dem Verweigerungsbias in ihrer Studie nicht auch ein erheblicher Nichterreichbarkeitsbias vorliegt.

Die Nichtbeachtung der Nichterreichbaren führt gleichzeitig zu einem weiteren analytischen Problem. Wird nämlich die Unterscheidung zwischen Unit-Nonresponse wegen Verweigerung und Unit-Nonresponse wegen Nichterreichbarkeit durch Nichtbeachtung einer der beiden Gruppen ignoriert, so kann ein wichtiger Wandel bei der Unit-Nonresponse in telefonischen Umfragen, wie ihn Steeh et al. (2001: 242) etwa für die USA aufgezeigt haben, gar nicht erkannt werden:

„[...] refusal rates are not decreasing or moderating because potential respondents are more willing to talk to survey interviewers over the telephone. Rather what we are seeing [...] may be a change in the character of nonresponse. Evidence of this is [...] clearly apparent in the metropolitan area where refusals significantly declined and noncontacts significantly increased.“

(vgl. auch: Collier/Bienstock 2007: 179) Den Verzicht auf eine Diskussion der Teilnahmeausfälle, die nach Ausfallarten differenziert, halten wir angesichts dieses Forschungsstandes für unplausibel und problematisch.

2 Vgl. die rudimentären Hinweise zur Stichprobenziehung von S/D auf S. 107 (Anm. 5); insbesondere der Hinweis auf das so genannte „Rösch-Design“.

6 Verhältnis von Panel- und Erstbefragten

S/D belegen die Qualität ihrer Verweigererstudie unter anderem auch mit dem Verweis auf die sehr hohe Ausschöpfungsquote (59 %, S/D: 111f.), die sie bei Ihrer Verweigererbefragung erreicht haben. Vor dem Hintergrund der Ausgangsfrage des Beitrages (*States or Traits?*) verliert diese Ausschöpfungsquote aber dadurch an Überzeugungskraft, dass sowohl Erst- als auch Panelbefragte in der Studie befragt wurden, ohne dass S/D die Leserschaft darüber aufklären, wie hoch der Anteil der Panel- und Erstbefragten in der Stichprobe ist. Das ist insofern von Bedeutung, als man bei den Panelbefragten davon ausgehen muss, dass sie weder allgemein noch bezogen auf diesen Studiengegenstand strukturelle Verweigerer sind, denn als Panelbefragte haben sie ja mindestens bereits einmal an genau dieser Umfrage teilgenommen. Gibt es also stabile „Traits“, – eine Frage, die mit der hier kritisierten Studie untersucht werden soll – so ist bei diesen Befragten davon auszugehen, dass tatsächlich vor allem situative Faktoren die Verweigerung beim Erstanruf erklären, während grundsätzlich eine Teilnahmebereitschaft vorhanden ist. Gleichzeitig ist aber von einer Konvertierbarkeit dieser Befragten auszugehen, die über der einer Zufallsstichprobe der Bevölkerung liegt, wie die Erstbefragten sie darstellen. Die leichtere Konvertierbarkeit der Panelbefragten mindert daher den Eindruck, den die hohe Ausschöpfung von 59 % hinterlassen soll. Die aufkommenden Zweifel hätten gemindert werden können, wenn S/D, wie oben bereits angedeutet, über den Anteil von Erst- und Panelbefragten berichtet hätten. Dies tun sie jedoch nur an einer Stelle, wenn nämlich berichtet wird, dass von den 139 Personen, die die Nachbefragung durch die Autoren verweigerten, 24 Panel- und 115 Erstbefragte waren. Diese Information ist insofern wertlos, als man ohne eine Information über das Ausgangsverhältnis nichts über unterschiedliche Teilnahmebereitschaften erfährt.

Den bereits gefallenem Begriff der Konversion aufgreifend möchten wir außerdem darauf verweisen, dass es sinnvoll gewesen wäre, die Nachbefragung als Wiederholungsbefragung mit Konversionscharakter zu interpretieren. Dies tut explizit Schräpler (2000) mit einer Verweigereranalyse der SOEP-Befragten. Die von Schräpler berichtete Konversionsquote durch Nachfrage bei Erstverweigerern lag bei 15 %. Interpretiert man die Verweigererbefragung als Konversionsversuch, dann verändert dies notwendigerweise auch die Erwartungen an die Befragten. Man versteht die Befragten bei dieser Herangehensweise als Personen mit unterschiedlicher Bereitschaft zur Befragungsteilnahme und eher nicht als eine Gruppe von Befragungsteilnehmern und eine Gruppe von Verweigerern. Mit Blick auf die S/D leitende Frage nach dem Stellenwert von States und Traits ist dies eine relevante Verschiebung der Perspektive. Außerdem ließen sich an dieses Verständnis kontinu-

ierlicher Unterschiede in der Teilnahmebereitschaft weitere Fragen anschließen, die leider in dem Artikel von S/D gar nicht diskutiert werden. Zu diesen Fragen gehört u. a., wie Befragte trotz einer ersten negativen Reaktion zu einer Umfrageteilnahme bewegt (konvertiert) werden können (vgl. dazu nochmals Schröpfer 2000).

Es sei an dieser Stelle erwähnt, dass die Autoren bei allen Analysen des Abschnittes 5 regelmäßig darauf verweisen, dass sich Panel- und Erstbefragte in ihrem Antwortverhalten nicht signifikant unterscheiden (S/D: 107). Damit wird zwar deutlich, dass S/D sich mit den Unterschieden beider Gruppen befasst haben, unverständlich bleibt aber, warum der Leserschaft die Details dieser Analyse wie der Verteilung von Panel- und Erstbefragten vorenthalten werden. Beides hätte den Ergebnissen höhere Glaubwürdigkeit verleihen können.

7 Fazit

Schnauber und Daschmann stellen in ihrem Beitrag die Frage, ob es stabile Verweigererdispositionen gebe, die systematisch Ausfälle bei (Telefon)surveys erzeugen und damit die Eignung solcher Umfragedaten für repräsentative Aussagen über Grundgesamtheiten in Frage stellen. Ihre Antwort lautet nein. Wie wir zeigen konnten, weist die Argumentation von S/D eine Reihe von Problemen auf. Zuvörderst sind es Probleme der Interpretation ihrer statistischen Ergebnisse, die die sehr weitreichende Schlussfolgerung aus der Zusammenfassung in Frage stellen. Im Weiteren ist es die Nichtbefassung mit dem Problem der Nichterreichbarkeit, die die weitreichenden Schlussfolgerungen als diskussionsbedürftig erscheinen lässt.

Ein genereller Nachweis auch nur der Nichtexistenz eines durch Verweigerung verursachten Non-Response-Bias liegt nach unserer Einschätzung mit dem Beitrag von Schnauber und Daschmann nicht vor. Unit-Nonresponse, sowohl durch Verweigerung als auch durch Nichterreichbarkeit, bleibt also bis auf Weiteres ein Problem, mit dem sich die Umfrageforschung jeweils studienspezifisch auseinandersetzen muss. Dies macht es u. a. erforderlich, dass Umfragedaten Datensätze für Nichtrespondenten aufweisen, die so viel Information über die Befragungssituation und den Interviewer enthalten wie möglich, um eine sinnvolle Analyse der Nonresponse zu ermöglichen (Schnell 2008).

Literatur

- Atrostic, B. K., N. Bates, G. Burt und A. Silberstein, 2001: Nonresponse in U.S. Government household surveys: consistent measures, recent trends, and new insights. *Journal of Official Statistics* 17: 209–226.
- Backhaus, K., B. Erichson, W. Plinke und R. Weiber, 2008: *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. Berlin/Heidelberg: Springer.
- Baur, N., 2006: Ausfallgründe bei zufallsgenerierten Telefonstichproben am Beispiel des Gabler-Häder-Designs. S. 159–183 in: F. Faulbaum und C. Wolf (Hg.): *Stichprobenqualität in Bevölkerungsumfragen*. Bonn: Informationszentrum Sozialwissenschaften.
- Behnke, J., N. Baur und N. Behnke, 2006: *Empirische Methoden der Politikwissenschaft*. Paderborn/München/Wien/Zürich: Schöningh.
- Collier, J. E. und C. C. Bienstock, 2007: An analysis of how nonresponse error is assessed in academic marketing research. *Marketing Theory* 7: 163–183.
- Dethlefsen, H.-A., 2000: Qualitätsmanagement in der CATI-Forschung. S. 49–64 in: V. Hüfken (Hg.): *Methoden in Telefonumfragen*. Opladen: Westdeutscher Verlag.
- Gabler, S., J. H.P. Hoffmeyer-Zlotnik und D. Krebs, 1994: *Gewichtung in der Umfragepraxis*. Opladen: VS Verlag für Sozialwissenschaften.
- Groves, R. M. und M. P. Couper 1998: *Nonresponse in household interview surveys*. New York: Wiley.
- Heyde, C. v. d., 2002: Das ADM-Telefonstichproben-Modell. S. 32–45 in: S. Gabler, S. Häder (Hg.): *Telefonstichproben. Methodische Innovationen und Anwendungen in Deutschland*. Münster/New York/München/Berlin: Waxmann.
- Kozak, M., 2009: What is strong correlation? *Teaching Statistics* 31: 85–86.
- Niemann, S. und T. Abel, 2000: Stichprobenselektion in einer telefonischen Bevölkerungsbefragung: ein Vergleich von Teilnehmern und Nichtteilnehmern im Berner Lebensstil-Panel. S. 105–122 in: V. Hüfken (Hg.): *Methoden in Telefonumfragen*. Opladen: Westdeutscher Verlag.
- Reuband, K.-H. und J. Blasius, 2000: Situative Bedingungen des Interviews, Kooperationsverhalten und Sozialprofil konvertierter Verweigerer. Ein Vergleich von telefonischen und face-to-face-Befragungen. S. 139–169 in: V. Hüfken (Hg.): *Methoden in Telefonumfragen*. Opladen: Westdeutscher Verlag.
- Schendera, C. F.G., 2008: *Regressionsanalyse mit SPSS*. München: Oldenbourg.
- Schnauber, A. und G. Daschmann, 2008: States or Traits? Was beeinflusst die Teilnahmebereitschaft an telefonischen Interviews? *Methoden – Daten – Analysen* 2: 97–123.
- Schnell, R., 2008: Antworten auf Nonresponse. Sozialwissenschaftlicher Informationsdienst. *Methoden und Instrumente der Sozialwissenschaften* 2008/1-1: 11–23. http://www.gesis.org/fileadmin/upload/dienstleistung/fachinformationen/servicepublikationen/sofid/Fachbeitraege/Methoden_2008-1-1.pdf (9.9.2009).
- Schnell, R., P. B. Hill und E. Esser, 2005: *Methoden der empirischen Sozialforschung*. München/Wien: Oldenbourg.
- Schnell, R., 1997: *Nonresponse in Bevölkerungsumfragen*. Opladen: Leske Budrich.
- Schnell, R., 1993: Die Homogenität sozialer Kategorien als Voraussetzung für „Repräsentativität“ und Gewichtungsverfahren. *Zeitschrift für Soziologie* 22: 16–32.
- Schräpler, J.-P., 2000: Was kann man am Beispiel des SOEP bezüglich Nonresponse lernen? *ZUMA-Nachrichten* 46: 117–149. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_46.pdf (9.9.2009).
- Steeh, C., N. Kirgis, B. Cannon und J. DeWitt, 2001: Are they really as bad as they seem? Nonresponse rates at the end of the Twentieth Century. *Journal of Official Statistics* 17: 227–247.
- Tuckel, P. und H. O'Neill, 2002: The vanishing respondent in telephone surveys. *Journal of Advertising Research* (September – Oktober): 26–48.

Weins, C., 2006: Multiple Imputation. S. 205–216 in: J. Behnke, T. Gschwend, D. Schindler und K.-U. Schnapp (Hg.): Methoden der Politikwissenschaft: Neuere qualitative und quantitative Analyseverfahren. Baden-Baden: Nomos.

Anschrift der Autoren

Prof. Dr. Kai-Uwe Schnapp
Olaf Bock
Universität Hamburg
Fakultät Wirtschafts- und
Sozialwissenschaften
Fachbereich Sozialwissenschaften
Institut für Politikwissenschaft
Allende-Platz 1
20146 Hamburg
kai-uwe.schnapp@wiso.uni-hamburg.de
olaf.bock@wiso.uni-hamburg.de