

Mobiltelefon- erfahrung und Antwortqualität bei Umfragen

Experience with Mobile Phones and Survey Response Quality

Michael Häder und Mike Kühne

Zusammenfassung

Telefonische Befragungen über das Mobilfunknetz sind inzwischen ein Bestandteil der sozialwissenschaftlichen Datengewinnung. Da es sich beim Mobilfunk jedoch noch um eine relativ junge Befragungstechnologie handelt, stellt sich die Frage, inwieweit die Routine im Umgang mit dem Handy als eine Determinante der Antwortqualität wirkt. Dieser Frage wurde im Rahmen des Projektes CELLA nachgegangen (Telefonbefragungen in der Allgemeinbevölkerung über das Mobilfunknetz 2009, GESIS Köln, Deutschland, DOI: 10.4232/1.4875). CELLA wurde durch die Deutsche Forschungsgemeinschaft gefördert und steht für telefonische Befragungen via Mobilfunk (*CELL* phone) und via Festnetz (*LAnd*line phone). Zunächst werden die Operationalisierung und die Bildung eines Index beschrieben, anhand dessen die Routine im Umgang mit dem Mobilfunk gemessen werden kann. Auf der Basis von sechs Kriterien der Antwortqualität wird dann nach den entsprechenden Zusammenhängen gesucht. Dabei stellt sich heraus, dass sich – entgegen unseren ursprünglichen Erwartungen – solche Zusammenhänge jedoch kaum nachweisen lassen. Die Befunde sind für den Einsatz von telefonischen Befragungen über das Mobilfunknetz insgesamt positiv zu bewerten. Sie

Abstract

Telephone surveys via mobile phones have become increasingly important for the social sciences. The use of mobile phones in interviews is still a novelty. Thus the question arises to what extent the respondents' experience with mobile phones influences the response quality. The German Research Foundation funded a CELLA project to investigate how to conduct mobile phone surveys. CELLA stands for *CELL* phone and *LAnd*line phone surveys. Technische Universität Dresden, Germany, and GESIS – Leibniz Institute for Social Sciences, Germany, cooperated for that purpose. Data presented in the following section were collected within the frame of CELLA (Telefonbefragungen in der Allgemeinbevölkerung über das Mobilfunknetz 2009, GESIS Köln, DOI: 10.4232/1.4875).

Firstly, an index for the experience in handling mobile phones is presented. This is followed by an examination of correlations between the index and six aspects of response quality. A significant influence of the experience was – against our expectations – not found. The results support the conduct of telephone surveys via mobile phones. Nevertheless, there are numerous determinants with stronger influence, which suggests that further research needs to be done.

verdienen es aber auch, bei der Diskussion um die Determinanten der Antwortqualität bei Umfragen allgemein berücksichtigt zu werden. So zeigt sich vor allem, dass von einer ganzen Reihe an Determinanten ein stärkerer Einfluss auf die Antwortqualität ausgeht als von der Routine.

1 Einleitung¹

Telefonische Befragungen in der Allgemeinbevölkerung, die auf Kontakte über das Mobilfunknetz verzichten, liefern inzwischen aufgrund von Coverage-Problemen nur noch verzerrte Ergebnisse. Verantwortlich dafür ist vor allem die Gruppe der Mobile-Onlys (Kuusela/Callegaro/Vehovar 2008; Blumberg/Luke 2008). Dabei handelt es sich um Personen, die nur noch über ein Mobiltelefon erreichbar sind, die also nicht mehr über einen Festnetzanschluss verfügen. Dies betrifft gegenwärtig etwa 11 % der Bevölkerung der Bundesrepublik. Vermutlich wird dieser Anteil in der nächsten Zeit noch ansteigen. Insbesondere ist zu beachten, dass sich dieser Personenkreis durch ganz spezielle sozio-demografische Charakteristika auszeichnet und nicht etwa einen Querschnitt der Gesellschaft darstellt (vgl. Graeske/Kunz 2009: 57ff.).

Die damit notwendig werdenden Befragungen über das Mobilfunknetz bringen jedoch für die Umfrageforschung auch methodische Probleme mit sich. Spätestens seit den Arbeiten von Dillman (vgl. 2000) ist bekannt, dass die Routine im Umgang mit dem Medium, über welches eine Befragung vermittelt wird, für die Qualität der Antworten mit verantwortlich ist. Entsprechende Effekte sind beispielsweise von ihm und seinen Kollegen bei Internetbefragungen nachgewiesen worden, die sie dann den Ergebnissen von postalischen und persönlich-mündlichen Befragungen gegenüber gestellt haben (vgl. Christian/Dillman/Smyth 2008: 257).

Auch bei der Kommunikation mittels Mobilfunk handelt es sich um eine noch relativ junge Technologie. Trotz der hohen Penetrationsrate – statistisch gesehen verfügt jeder Haushalt in der Bundesrepublik bereits über mehr als eine Mobilfunknummer (vgl. Häder/Gabler/Heckel 2009) – kann noch nicht davon ausgegangen werden, dass der Umgang mit diesen Geräten, etwa im Vergleich mit Festnetztelefonen, bereits bei allen Nutzern zur Routine geworden ist.

1 Die Autoren bedanken sich bei den beiden anonymen Gutachtern für die zahlreichen Hinweise zu ihrem Manuskript. Sie waren uns eine große Hilfe bei der Fertigstellung des vorliegenden Beitrages.

Der vorliegende Text wendet sich anhand empirischer Daten deshalb der Frage zu, ob ein und – falls gegeben – welcher Einfluss von der Routine im Umgang mit den Mobilfunkgeräten auf die Qualität der Antworten bei einer sozialwissenschaftlichen Befragung via Mobiltelefon ausgeht.

Bei der für diese Analysen benutzten Datenbasis handelte es sich um die Studie CELLA 1² (*Cell-Phone* und *Land-Line*). Es wurden hier insgesamt 1.009 zufällig ausgewählte Personen über das Festnetz und 1.162 Personen über den Mobilfunk mit den gleichen Instrumenten befragt. Grundlage der Stichprobenziehung für die Befragung über das Festnetz war das Gabler-Häder-Design (vgl. Gabler/Häder 2009). Für die Auswahl der Teilnehmerinnen und Teilnehmer für die Mobilfunkbefragung wurde ein modifiziertes RDD-Verfahren verwendet (vgl. Häder/Gabler/Heckel 2009: 21ff.).

Im folgenden Abschnitt wird, nachdem einige theoretische Überlegungen zur Erklärung des Antwortverhaltens dargestellt wurden, die Operationalisierung der Routine im Umgang mit dem Mobiltelefon und der drauf aufbauende Index präsentiert. Dieser drückt aus, inwieweit die Zielpersonen im Umgang mit ihrem Handy vertraut sind. Im Anschluss werden dann anhand des genannten Datensatzes (vgl. Häder/Häder 2009: 269ff.) die verschiedenen Kriterien für die Antwortqualität mit dem zuvor gebildeten Routineindex in Beziehung gesetzt. Abschließend wird versucht, die empirischen Befunde auch theoretisch einzuordnen.

2 Kognitionspsychologische Modelle der Antwortgenerierung: Implikationen für die Antwortqualität

Kognitive Modelle des Antwortprozesses (vgl. z. B. Sudman/Bradburn/Schwarz 1996; Tourangeau/Rips/Rasinski 2000) stellen eine hilfreiche Basis zur Erklärung des Antwortverhaltens bereit. Danach sind erstens das Verständnis der Frage, zweitens der Abruf von Informationen aus dem Gedächtnis, drittens die Bildung einer Antwort sowie viertens das Formulieren der Antwort die wesentlichen Schritte, die eine Person bei einer Befragung zu absolvieren hat. Insbesondere die Theorien paralleler

2 Der Datensatz der CELLA 1 Untersuchung ist beim Datenarchiv der GESIS – Leibniz Institut für Sozialwissenschaften – erhältlich (DOI: 10.4232/1.4875). Eine ausführliche Darstellung der gesamten Untersuchung einschließlich des Fragebogens findet sich bei Häder/Häder (2009). Die in diesem Beitrag verwendeten Variablenbezeichnungen sind mit denen im Datensatz identisch. Die Studie CELLA 2 wird 2010 durchgeführt. Die Untersuchung CELLA 1 zu Telefonbefragungen über das Mobilfunknetz wurde von der DFG gefördert. Die Förderung für CELLA 2 erfolgt ebenfalls von dort.

Verarbeitungswege haben der empirischen Umfrageforschung für die Darstellung zahlreicher empirischer Phänomene den theoretischen Rahmen liefern können.

Aufgrund der Komplexität des Antwortprozesses erscheint es jedoch fraglich, inwieweit die Befragten tatsächlich, wie in den hier ausgearbeiteten kognitionspsychologischen Pfadmodellen beschrieben (vgl. Strack 1994: 53; Sudman/Bradburn/Schwarz 1996: 56; Tourangeau/Rips/Rasinski 2000: 8; Jobe/Herrmann 1996), den kompletten Prozess der Antwortgenerierung durchlaufen (vgl. Tourangeau/Rips/Rasinski 2000: 354). Es ist vielmehr davon auszugehen, dass es bei ungünstigen Bedingungen während des Interviews – beispielsweise bei kognitiv besonders anspruchsvollen Sachverhalten, bei einer niedrigen Salienz des in der Frage aufgegriffenen Gegenstandes oder, wie an dieser Stelle angenommen, bei mangelnder Routine im Umgang mit dem Kommunikationsmedium, über welches die Befragung geführt wird – zum Einsatz von Heuristiken bei der Antwortgenerierung kommt (vgl. Strack/Martin 1988). Diesem Modell weiter folgend liegt der Antwort dann keine kognitiv anspruchsvolle Urteilsbildung zugrunde, welche auf den individuellen bisherigen Wissensbeständen zum erfragten Sachverhalt basiert. Vielmehr entscheiden sich die Befragten bei ihrer Antwort für eine vermeintlich einfache kognitive Lösung. Wenn überhaupt eine Antwort übermittelt wird, ist diese relativ spontan. Problematisch ist, dass damit Beeinträchtigungen der Datenqualität einhergehen, da sowohl eine erhöhte Tendenz zu Item-Nonresponse als auch zu einer zufälligen Entscheidung für eine Antwortkategorie feststellbar sind.

Krosnick und Alwin (1987; vgl. auch Strack/Martin 1988) entwickelten ein Modell paralleler Verarbeitungswege, das über den Spezialfall der Meinungsfragen hinaus auf alle Phasen des Antwortprozesses angewendet werden kann. Dabei wird insbesondere die Motivation der Befragten in den Mittelpunkt gestellt; und es wird davon ausgegangen, dass – beispielsweise wiederum aufgrund fehlender Routine im Umgang mit dem Mobiltelefon – weniger motivierte Befragte die kognitiven Verarbeitungsprozesse nur oberflächlich durchführen (Weak Satisficing) und einige diese sogar vollständig überspringen (Strong Satisficing). Sowohl das Strong- als auch das Weak Satisficing führen so zu Reihenfolgeeffekten (vgl. Payne 1949; Hyman/Sheatsley 1950; Dillman et al. 1995; für einen deutschsprachigen Überblick: Faulbaum/Prüfer/Rexroth 2009).

In zahlreichen Studien haben Umfrageforscher eine Vielzahl von weiteren Effekten nachgewiesen, die infolge der genannten Satisficing-Strategien hervorgebracht werden und die damit zu systematischen Verzerrungen des wahren Wertes führen. Dazu zählen beispielsweise auch Skaleneffekte (vgl. Stadtler 1983; Schwarz/Hippler/Deutsch/Strack 1985; Bishop 1987; Böcker 1988) und Kontexteffekte (vgl. Cantril 1944, zitiert nach Strack 1994; Schuman/Presser 1981: 47; für einen Über-

blick: Schuman 1992). Je stärker die hier aufgezählten Effekte auftreten, desto geringer ist die Qualität der Antwort. Daraus lassen sich die folgenden Aspekte als Grundlage der sich anschließenden Analysen ableiten. Basierend auf den Annahmen der kognitionspsychologischen Modelle der Antwortgenerierung lässt sich eine hohe Antwortqualität erwarten, wenn die Befragten alle vier Phasen des Antwortprozesses ungestört durchlaufen. Im Gegensatz dazu sollte eine mangelnde Routine im Umgang mit dem Mobilfunk zu Ablenkung im Antwortprozess und damit zu weniger elaborierten Antworten und damit zu einer geringeren Antwortqualität führen.

3 Operationalisierung der unabhängigen Variablen

Zwei Befragungen, die sich auf die gleiche Grundgesamtheit beziehen und das gleiche Instrument verwenden, müssen nicht zwingend zu den gleichen Befunden führen. Insbesondere, wenn verschiedene Erhebungsmodes verwendet werden, kann es zu Unterschieden im Antwortverhalten kommen (Häder/Kühne 2009a: 165ff.). Ursachen dafür liegen unter anderem in modespezifischen Eigenschaften der Erhebung. Für eine Befragung über den Mobilfunk lassen sich einige Besonderheiten vermuten (Häder/Kühne 2009a: 165). Schwerpunkt der folgenden Analysen wird die Routine im Umgang mit dem Mobiltelefon sein.

Routine im Umgang mit dem Mobiltelefon

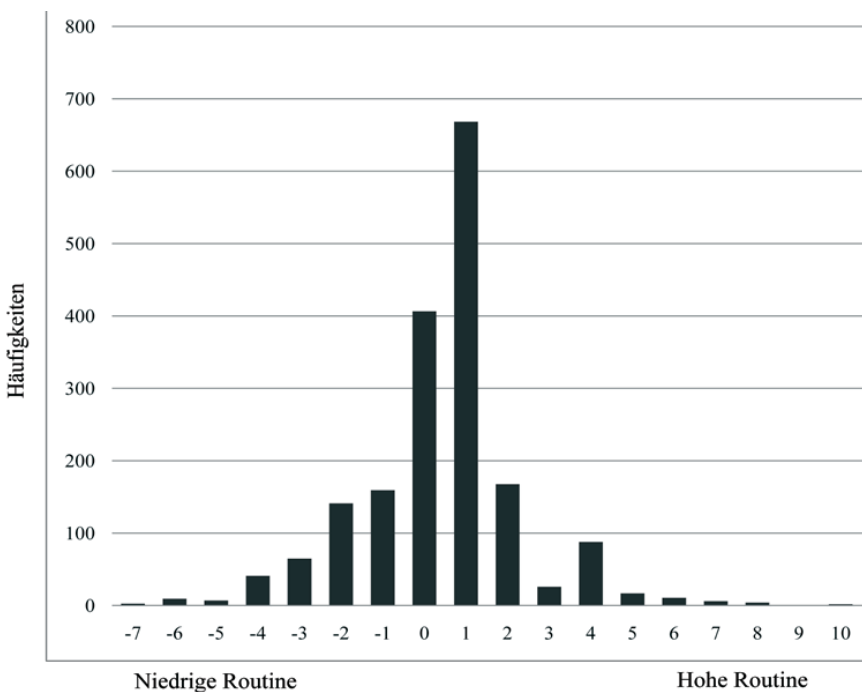
Die Erfahrung im Umgang mit Mobiltelefonen lässt sich nur schwer anhand einer einzelnen Frage messen. Es ist aber davon auszugehen, dass ein Anstieg in der Intensität der Telefonnutzung zu einem Zuwachs an Erfahrung im Umgang führt. Die Intensität der Handynutzung wurde anhand der folgenden Indikatoren³ gemessen:

- Anzahl der Handynummern, über die die Zielperson privat erreichbar ist (v4),
- Häufigkeit, mit der auf dem eigenen Handy Anrufe eingehen (v9),
- Häufigkeit, mit der das Handy eingeschaltet ist (v8_1) und
- Häufigkeit, mit der das Handy auch unterwegs mitgenommen wird (v11).

3 Es wurde auch getestet, ob die Einbeziehung weiterer Variablen in die Berechnung des Index zu einer Verbesserung von dessen Güte führt. Dabei handelte es sich um Fragen danach, ob eingehende Anrufe verpasst werden (v10), um die Anzahl der Rufnummern, die auf dem eigenen Handy gespeichert sind (v16) und um eine Frage danach, ob Anrufe in der Öffentlichkeit als peinlich empfunden werden (v17). Da dies nicht der Fall war, wurde auf die Berücksichtigung dieser Indikatoren verzichtet.

Diese Variablen wurden standardisiert und zu einem additiven Index zusammengefasst, bei dem ansteigende Werte eine zunehmende Routine repräsentieren. Der Index weist eine ausreichende interne Konsistenz (Cronbachs Alpha = 0,77) auf und ist eindimensional. Aus Abbildung 1 ist die Verteilung des Index zu entnehmen. Dabei ist zu erkennen, dass die Routine im Umgang mit dem Mobilfunkgerät in der Stichprobe durchaus unterschiedlich vorhanden ist und nicht etwa nur Personen mit einer ausgeprägten Technikaffinität befragt worden sind.

Abbildung 1 Häufigkeitsverteilung des Routine-Index



Eine erste nähere Betrachtung macht deutlich, dass es – wie auch nicht anders zu erwarten war – einen Zusammenhang des Grades an Routine mit dem Alter der Befragten gibt. Der Korrelationskoeffizient nach Pearson beträgt $r=-0,39$ ($p<0,001$). Jüngere Personen besitzen damit eine höhere Routine im Umgang mit dem Handy. Auch ein klarer Geschlechterunterschied ist anhand der Daten zu verzeichnen: Männer verfügen über eine größere Routine als Frauen ($M_{\text{Männer}} = 3,30 > M_{\text{Frauen}} = 3,13$; $p<0,001$).

Unsere Bemühungen gehen dahin, wahre Effekte, die von der Routine auf das Antwortverhalten ausgehen, zu ermitteln. Damit sind vor allem multivariate Auswertungsverfahren erforderlich, um einen möglichen Einfluss weiterer Drittvariablen kontrollieren zu können. Die folgenden Analysen beziehen sich ausschließlich auf die Befragten ($n=1.162$), die in der CELLA 1 Studie über ein Mobiltelefon kontaktiert und befragt wurden.

Kontrollvariablen

Personen können am Mobilfunk in verschiedenen sozialen Situationen und Umgebungen erreicht werden. Diese Umgebungen können sich in einzelnen Aspekten (z. B. Grad der Privatheit) erheblich unterscheiden. In der Befragung wurde deshalb auch erhoben, wo sich die befragten Personen zum Zeitpunkt des Interviews befanden. Der größte Teil der Befragten (63,2 %) wurde auch über den Mobilfunk zu Hause befragt.⁴

Bei der Auswertung wird der Ort des Interviews dichotomisiert verarbeitet. Der Wert 0 besagt, dass die Befragung zu Hause, und der Wert 1, dass die Befragung an einem anderen Ort stattfand.

Aufgrund der Heterogenität der Orte, an denen die Befragten sich befanden, wurde auch erhoben, ob zum Zeitpunkt des Interviews Dritte anwesend waren. Ist der Befragte zum Zeitpunkt des Interviews nicht allein, vergrößert sich das Publikum in der Interviewsituation. Die Wirkung der Anwesenheit Dritter auf den Antwortprozess ist allerdings nicht eindeutig. Bei günstigen Bedingungen – hierzu zählt beispielsweise auch eine hohe Routine im Umgang mit dem betreffenden Kommunikationsmedium – können damit die Anstrengungen, eine sozial erwünschte Antwort zu geben, intensiver ausfallen als in einem ungünstigen Setting.

Für die Auswertung wird die Anwesenheit dritter Personen während des Interviews ebenfalls dichotomisiert erhoben. Hier besagt der Wert 0, dass keine dritte Person während des Interviews anwesend war. Der Wert 1 bringt zum Ausdruck, dass sich zumindest zeitweise ein Dritter während des Interviews in unmittelbarer Nähe aufhielt.

Um potentielle Mediatoreffekte der Routine auf die Anwesenheit Dritter und auf den Ort des Interviews zu kontrollieren, wurden in den multivariaten Analysen diese Interaktionsterme mit einbezogen. Zusätzlich fanden die soziodemografischen Variablen Alter (in Jahren) Bildung (in Jahren) und das Geschlecht Eingang in die Modelle.

4 Aufgrund der Heterogenität der anderen Antworten wird diese Variable dichotom (zu Hause ja/nein) in den Analysen berücksichtigt.

4 Indikatoren der Antwortqualität

In der Umfrageforschung lassen sich vier zentrale Dimensionen zur Beurteilung der Qualität von Umfragen aufzeigen. Dabei handelt es sich – unter Verwendung der inzwischen gängigen englischen Begriffe – um: Sampling, Coverage, Response und Measurement (de Leeuw et al. 2008: 2f.). In diesem Aufsatz stehen das Antwortverhalten und damit insbesondere der Bereich der Messfehler im Vordergrund.

Zwei Gruppen von Indikatoren wurden für die Ermittlung der Qualität des Antwortverhaltens bzw. zur Aufdeckung von Messfehlern verwendet. Erstens globale Indikatoren, die die Qualität verschiedener Schritte der Antwortfindung (vgl. Abschnitt 2) zugleich abbilden, und zweitens Indikatoren, die die Qualität eines bestimmten Schrittes der Antwortfindung wiedergeben.⁵ Zur ersten Gruppe von Indikatoren zählen:

- Pseudo-Opinions: Eine Reaktion auf fiktive Fragegegenstände könnte sowohl auf ein mangelndes Verständnis der Frage als auch auf einen mangelhaften Informationsabruf oder auf sozial erwünschtes Antwortverhalten zurückgeführt werden (Kühne/Böhme 2006). Als eine Folge geringer Routine im Umgang mit der Kommunikationstechnik lässt sich deshalb ein stärkeres Auftreten von Pseudo-Opinions vermuten.
- Item-Nonresponses: Ursachen für die Nichtbeantwortung einzelner Fragen können in allen vier Phasen des kognitionspsychologischen Antwortmodells vermutet werden. Mangelnde Routine sollte insbesondere zu einer höheren Wahrscheinlichkeit von Satisficing-Strategien führen, die wiederum Item-Nonresponse zur Folge haben können (Krosnick 2000). Substanzielle Antworten werden im umgekehrten Fall aufgrund eines höheren kognitiven Einsatzes gegeben, sodass seltener Satisficing-Strategien genutzt werden (Shoemaker/Eichholz/Skewes 2002: 196).
- Antwortstabilität: Die Ursache für instabile individuelle Einstellungen werden in der Umfrageforschung kontrovers diskutiert (u. a. Converse 1964, 1970; Kriesi 2004: 242; Zaller 1992; Alvarez/Brehm 1995, 2002). Allerdings lässt sich vermuten, dass instabile Einstellungen auf Ablenkungen im Antwortprozess zurückführbar sind (Kühne/Böhme 2006: 49f.). Insofern sollte eine geringere Routine zu einem Anstieg instabiler Antworten führen.

Zur zweiten Gruppe von Indikatoren zählen zunächst solche, die ganz konkret auf den Informationsabruf für die Urteilsbildung zielen:

5 In CELLA 1 wurden zahlreiche weitere Indikatoren untersucht, die an dieser Stelle nicht präsentiert werden können. Unter anderem wurden der Zeitpunkt des Abbruchs und die Schwierigkeit der Beantwortung der Fragestellung analysiert (Häder/Häder 2009).

- **Erinnerungsleistung:** Die Erinnerungsleistung kann als Indiz dafür gelten, wann die Informationssuche – der Schritt zwei – für die Antwortfindung abgebrochen wurde bzw. mit welcher Intensität diese überhaupt geführt wurde (Häder/Kühne 2009b: 225ff.). Fehlende Routine im Umgang mit dem Kommunikationskanal kann bei den Befragten zu geringeren Erinnerungsleistungen führen. Es ist bekannt, dass für die Findung einer Antwort nicht alle zur Verfügung stehenden Informationen aus dem Gedächtnis abgerufen werden, sondern dass dieser Prozess abgebrochen wird, sobald die Zielperson meint, genügend erinnert zu haben. Dieser Abbruch erfolgt vermutlich um so eher, je geringer der Grad an Routine ist.
- **Reihenfolgeeffekte:** Treten in einer Befragung Reihenfolgeeffekte auf, kann das unter anderem auf Unterschiede bei dem Abruf von Informationen für die Antwortgenerierung hindeuten (Häder/Kühne 2009c: 207ff.). Bei fehlender Routine sollten systematische Effekte durch die Änderung der Fragereihenfolge stärker auftreten. Bei einem höheren Grad an Routine können sich die Befragten verstärkt auf ihre Antwort konzentrieren und sind gegenüber den genannten Einflüssen resistenter.

Die CELLA 1 Studie bietet damit zahlreiche Möglichkeiten, um den Einfluss der Routine auf die Qualität des Antwortverhaltens analog zu den formulierten Vermutungen empirisch zu überprüfen. Dies macht einen besonderen Wert der Studie aus.

5 Ergebnisse

5.1 Globale Indikatoren der Antwortqualität

Pseudo-Opinions

Pseudo-Opinions stellen einen beliebten Indikator der Antwortqualität dar (Bishop/Tuchfarber/Oldendick 1986; Reuband 2000; Mayerl/Urban 2008; Krosnick et al. 2002; Kühne/Böhme 2006). CELLA 1 enthält eine Frage zu einem Sachverhalt (konkret: die Bewertung eines Vorschlags des Europäischen Parlaments zum Ausbau eines gemeinsamen, europaweiten Mobilfunknetzes, tp3), den es nicht gegeben hat. Eine positive Antwort auf eine solche Frage deutet auf eine geringere Antwortqualität hin.

Überraschend ist zunächst, dass die Befragten, die angaben, diesen Vorschlag zu kennen ($n=143$), über eine etwas höhere Routine verfügen als jene ($n=772$), die dies verneinen ($M_{\text{bekannt}} = 0,35 > M_{\text{nicht bekannt}} = -0,04$; $p < 0,05$; gewichteter Datensatz).

Diejenigen Befragten, die auf das Pseudo-Opinion reagieren, haben also eine etwas größere Routine im Umgang mit dem Handy. Weiteren Aufschluss kann das Regressionsmodell liefern (vgl. Tabelle 1). In dieses Modell werden auch zwei Interaktionsterme aufgenommen. Der eine betrifft die Routine und die Anwesenheit Dritter (*Routine x Anwesenheit Dritter*), der andere die Routine und den Ort (*Routine x Ort des Interviews*), an dem das Interview stattfindet. Damit soll ein moderierender Einfluss der Routine kontrolliert werden. Es ist zu erkennen, dass auch unter Kontrolle der einbezogenen Größen der Einfluss der Routine auf das Antwortverhalten erhalten bleibt, während von den Interaktionstermen kein signifikanter Effekt ausgeht.

Tabelle 1 Ergebnisse einer binär logistischen Regression zur Erklärung des Antwortverhaltens bei einer Pseudo-Opinionfrage, Effektkoeffizienten

Konstante	8,50***
Routine	0,79**
Anwesenheit Dritter (Ref. Kat. = niemand anwesend)	1,14
Ort des Interviews (Ref. Kat. = zu Hause)	1,06
Alter	0,99
Geschlecht (Ref. Kat. = männlich)	1,48
Bildung	0,95
Routine x Anwesenheit Dritter	1,10
Routine x Ort des Interviews	1,21

n = 949, *Pseudo-R*² = 0,02, * = 0,05, ** = 0,01, *** = 0,001;
 Abhängige Variable: Pseudo-Opinion, Ref. Kat. = Reaktion auf Pseudo-Opinion

Item-Nonresponse

Die Stärke des Item-Nonresponse kann ebenfalls als Qualitätskriterium für das Antwortverhalten angesehen werden. Zu vermuten ist, dass bei größerer Routine für die Zielpersonen die Antwortfindung weniger kognitiv anspruchsvoll ist und ihnen der Abruf von Informationen aus dem Gedächtnis dadurch erleichtert wird. Weiter ist naheliegend, dass bei der Antwortfindung weniger Heuristiken eingesetzt werden und dass es damit zu weniger Item-Nonresponse kommt.

Für die Ermittlung des Ausmaßes an Item-Nonresponse sind insgesamt 50 Variablen der Studie CELLA 1 herangezogen worden, darunter beispielsweise auch

die Einkommensfrage.⁶ Bereits die bivariaten Analysen zwischen der Häufigkeit des Nonresponse und dem Routineindex verweisen auf einen fehlenden Zusammenhang. Der Korrelationskoeffizient nach Pearson ist minimal und nicht signifikant ($r=-0,013$; $p=0,67$; $n=1.123$, gewichteter Datensatz). Damit tritt der erwartete Zusammenhang nicht auf.

Eine detaillierte Aussage soll dabei wieder ein Regressionsmodell liefern (vgl. Tabelle 2). Ganz offenbar stützen die multivariaten Analysen die bivariaten Ergebnisse. Von den einfachen Kontrollvariablen hat einzig das Geschlecht einen Einfluss auf den Item-Nonresponse. Männer verweigern häufiger bei einzelnen Fragen die Antwort.

Tabelle 2 Ergebnis einer linearen Regression zur Erklärung der Häufigkeit des Item-Nonresponse, unstandardisierte und [standardisierte] Regressionskoeffizienten

Konstante	1,57***
Routine	-0,06 [-0,09]
Anwesenheit Dritter (Ref. Kat. = niemand anwesend)	0,03
Ort des Interviews (Ref. Kat. = zu Hause)	-0,08
Alter	0,01 [0,07]
Geschlecht (Ref. Kat. = männlich)	-0,40**
Bildung	0,02 [0,05]
Routine x Anwesenheit Dritter	0,10 [0,10]
Routine x Ort des Interviews	-0,13 [-0,12]

$n = 368$, $R^2 = 0,06$, * $=0,05$, ** $=0,01$, *** $=0,001$;

Abhängige Variable: Anzahl des Item-Nonresponse von 0 bis 14

Antwortstabilität

Wird innerhalb einer Befragung die Bewertung eines Sachverhalts wiederholt, so sollten dabei möglichst identische Befunde ermittelt werden. Ist dies nicht der Fall, so liegt der Schluss nahe, dass von den Zielpersonen lediglich Heuristiken bei der Antwortfindung benutzt werden und diese damit instabil sind. Die Antwort ist dann mehr oder weniger zufällig zustande gekommen. Dies liegt nahe, da bei einer hohen Qualität des Antwortverhaltens ein einmal gebildetes Urteil wiederholt

6 Es handelt sich konkret um die Folgenden: v2, v13, v8, v9, v10, v11, v13, f2a, f2b, tp1, tp2, tp3, tp4, v15_1, v15_2, v15_3, v15_4, v15_5, v17, pt5, v17a, v18_1, v18_2, v18_3, v18_4, v18_5, v18_6, v18_7, v18_8, v18_9, v18-1014, v19a, s1, s1_1, f1, f3_4, f3_5, f3_6, f3_7, f3_8, f3_9, s3, t4, s4, s4_1, s5, s5_2, s7, v22 und v28.

werden sollte. Werden die entsprechenden Fragen jedoch relativ willkürlich beantwortet, so liegt eine nur geringe Qualität des Antwortverhaltens vor.

In der CELLA 1 Studie waren zwei Fragen enthalten, die auf die Bewertung eines nahezu identischen Sachverhalts zielen. Relativ zu Beginn wurde gefragt (tp1), ob der Ausbau der Mobilfunknetze die Lebensqualität eher positiv oder eher negativ beeinflussen wird. Gegen Ende der Befragung wurde Auskunft darüber eingeholt (tp5), ob nach Meinung der Befragten beim Ausbau der Mobilfunknetze die Vorteile oder die Nachteile überwiegen (vgl. dazu auch Kühne/Häder 2009: 217ff.). Damit kann nun untersucht werden, ob eine höhere Routine im Umgang mit dem Handy zu stabileren Antworten führt.⁷ Das Ergebnis kann aus Tabelle 3 entnommen werden. Die Routine im Umgang mit dem Handy sowie auch die Interaktionsterme haben keinen Einfluss auf diesen Aspekt der Qualität des Antwortverhaltens. Die Kontrollvariablen Geschlecht und Bildung besitzen einen Einfluss. Frauen und Personen mit niedriger Bildung neigen demnach eher zu instabilem Antwortverhalten.

Tabelle 3 Ergebnisse einer binär logistischen Regression zur Erklärung der Antwortstabilität, Effektkoeffizienten

Konstante	3,78**
Routine	0,98
Anwesenheit Dritter (Ref. Kat. = niemand anwesend)	0,86
Ort des Interviews (Ref. Kat. = zu Hause)	1,08
Alter	1,00
Geschlecht (Ref. Kat. = männlich)	0,67*
Bildung	0,95*
Routine x Anwesenheit Dritter	1,03
Routine x Ort des Interviews	1,10

n = 891, *Pseudo-R*² = 0,02, * = 0,05, ** = 0,01, *** = 0,001;

Abhängige Variable: Antwortstabilität, Ref. Kat. = keine Antwortstabilität

7 Untersuchungen zur Stabilität von Einstellungen beruhen zum überwiegenden Teil auf Längsschnittdaten (Converse 1964; Schuman/Presser 1981; Craig/Kane/Martinez 2000). Es konnte aber bereits gezeigt werden, dass Antwortstabilität auch in Querschnittsdesigns gemessen werden kann (Slaby 1998; Kühne/Böhme 2006).

5.2 Indikatoren zum Informationsabruf und zum Editieren der Antworten in Richtung der sozialen Erwünschtheit

Erinnerungsleistung

Um die Erinnerungsleistungen der Zielpersonen zu ermitteln, wurde ein an die Methode des Paraphrasing (vgl. Driscoll 2007) angelehntes Vorgehen gewählt. Das Frageprogramm der CELLA 1 Studie enthält einen entsprechenden Indikator (v18_10). Nach der Abfrage einer Reihe von Sachverhalten – insgesamt wurden neun Einzelfragen gestellt – bestand die Aufgabe für die Zielperson darin, möglichst viele der zuvor von ihr bewerteten Dinge aus dem Gedächtnis zu wiederholen. Die Anzahl der erinnerten Sachverhalte gibt eine Auskunft über die Intensität der Informationssuche und erfüllt damit eine Indikatorfunktion für die Antwortqualität bei den übrigen Fragen.

Die Korrelation der Anzahl der erinnerten Gegenstände mit dem Routineindex zeigt keinen Zusammenhang ($r=-0,01$, $p=0,72$; $n=1.132$, gewichteter Datensatz). Das Ergebnis der Schätzung einer linearen Regression zur Erklärung der Anzahl von erinnerten Gegenständen aus den vorangegangenen neun Fragen zeigt Tabelle 4. Die Bildung und das Alter stellen sich, wie zu erwarten war, als Bestimmungsgrößen für die Erinnerungsleistung heraus. Auch vom Geschlecht geht ein signifikanter Effekt aus. Die Routine im Umgang mit dem Mobiltelefon sowie die beiden Interaktionsterme haben dagegen ein weiteres Mal keinen Einfluss.

Tabelle 4 Ergebnis einer linearen Regression zur Erklärung der Erinnerungsleistung, unstandardisierte und [standardisierte] Regressionskoeffizienten

Konstante	3,54***
Routine	-0,02 [-0,02]
Anwesenheit Dritter (Ref. Kat. = niemand anwesend)	0,04
Ort des Interviews (Ref. Kat. = zu Hause)	-0,20
Alter	-0,02 [-0,19]***
Geschlecht (Ref. Kat. = männlich)	0,33**
Bildung	0,09 [0,16]***
Routine x Anwesenheit Dritter	-0,06 [-0,04]
Routine x Ort des Interviews	0,02 [0,01]

$n = 1.084$, $R^2 = 0,06$, * = 0,05, ** = 0,01, *** = 0,001;
Abhängige Variable: Anzahl des erinnerten Items von 0 bis 9

Fragereihenfolge

Weitere deutliche Hinweise auf die Qualität des Antwortverhaltens liefert die Analyse von Fragereihenfolgeeffekten. Dabei liegt es nahe, dass mit steigender Größe dieses Einflusses die Qualität der Antworten sinkt – oder mit anderen Worten: Je mehr sich die Befragten bei ihren Antworten von der Reihenfolge der Fragen leiten lassen, desto weniger werden eigene Informationen aus dem Gedächtnis für die Antwortfindung aktiviert, desto weiter entfernt sind diese Angaben vom wahren Wert usw.

In unserem Zusammenhang ist zu analysieren, ob Menschen mit einer größeren Routine im Umgang mit dem Mobiltelefon resistenter gegenüber solchen Fragereihenfolgeeffekten sind. Schließlich wird angenommen, dass sich jene Befragten aufgrund ihrer Sicherheit im Umgang mit dem Medium Handy kognitiv stärker auf den Informationsabruf konzentrieren können als andere Personen.

In der Studie CELLA 1 war ein Split ballot enthalten: Einem zufällig rekrutierten Teil der Befragten wurde eine Frage zum allgemeinen Lebensglück (v17)⁸ vor und dem anderen Teil nach einer speziellen Indikatorenatterie zur Verzichtbarkeit bzw. Unverzichtbarkeit von materiellen Dingen (v18)⁹ gestellt. Mithilfe dieses Designs ist es gelungen, einen Reihenfolgeeffekt zu erzeugen. So kann im Weiteren festgestellt werden, ob die Stärke dieses Effekts im Zusammenhang steht mit der Routine, die die Zielpersonen im Umgang mit dem Mobilfunkgerät besitzen. Tabelle 5 zeigt das Ergebnis der Schätzung eines entsprechenden linearen Regressionsmodells.

Der Befund ist sehr aufschlussreich: So kann ein deutlicher Reihenfolgeeffekt ausgemacht werden (vgl. die Variable *Split Reihenfolge* in Tabelle 5). Auch die anderen zur Kontrolle in das Modell aufgenommenen Größen besitzen – wie zu erwarten war – teilweise eine starke erklärende Wirkung. Ohne Einfluss bleibt dagegen der Index zur Abbildung der Routine im Umgang mit dem Handy. Beachtung verdienen auch die drei in dieses Modell aufgenommenen Interaktionsvariablen. Neben den bereits bekannten wurde zusätzlich ein Term integriert (Variable *Routine x Split Reihenfolge*), der die Wirkung von Routine sowie den Split berücksichtigt. Der an dieser Stelle wiederum fehlende Effekt belegt, dass die Routine ein weiteres Mal ohne Einfluss auf die Antwortqualität bleibt.

8 "Wenn Sie einmal Ihr Leben betrachten, was würden Sie sagen, wie glücklich oder unglücklich sind Sie alles in allem?"

9 "Nun einmal zu einem anderen Thema. Es gibt verschiedene Meinungen darüber, welche Dinge man in Deutschland zum Leben braucht. Ich lese Ihnen im Folgenden einige Dinge vor und Sie sagen mir jeweils, was für Sie verzichtbar und was unverzichtbar ist. Markenartikel kaufen zu können, halten Sie für verzichtbar oder unverzichtbar. Ein neues Auto kaufen zu können, halten Sie für verzichtbar oder unverzichtbar. Über ein neues Handy zu verfügen, halten Sie für verzichtbar oder unverzichtbar", usw.

Tabelle 5 Ergebnis einer linearen Regression zur Erklärung des Antwortverhaltens bei der Frage nach dem Lebensglück, unstandardisierte und [standardisierte] Regressionskoeffizienten

Konstante	1,95***
Routine	-0,04 [-0,10]
Anwesenheit Dritter (Ref. Kat. = niemand anwesend)	-0,13 [-0,08]*
Ort des Interviews (Ref. Kat. = zu Hause)	-0,04
Alter	0,01 [0,10]**
Geschlecht (Ref. Kat. = männlich)	-0,01
Bildung	-0,02 [-0,07]*
Routine x Anwesenheit Dritter	0,01 [0,01]
Routine x Ort des Interviews	0,02 [0,02]
Split Reihenfolge	0,10 [0,06]*
Routine x Split Reihenfolge	0,01 [0,04]

$n = 1.051$, $R^2 = 0,03$, * $=0,05$, ** $=0,01$, *** $=0,001$;

Abhängige Variable: Einschätzung des Lebensglücks 1 = voll und ganz glücklich bis 5 = überhaupt nicht glücklich

Frageformulierung

Einen guten Hinweis auf die Qualität des Antwortverhaltens können auch Frageformulierungseffekte liefern. Sicher ist dabei: Je stärker von unterschiedlichen Frageformulierungen die Effekte auf das Antwortverhalten – bei identischem Inhalt der Frage – ausfallen, desto geringer wird die Antwortqualität. In der CELLA 1 Studie wurde ein in der Umfrageforschung seit längerem bekannter Test (vgl. Rugg 1941) modifiziert. Reuband (2003: 87ff.) verweist bereits auf eine ganze Reihe an Replikationen dieses Experiments (vgl. Schuman/Presser 1981; Hippler/Schwarz 1986; Waterplas et al. 1988; Glendall/Hoek 1990; Loosveldt 1997 und Holleman 2000). Die Zielpersonen sollten bei CELLA 1 angeben (f2), ob Handy-Sendemasten auf Schulgebäuden „erlaubt“ oder „verboten“ (Splitvariante a) bzw. „erlaubt“ oder „nicht erlaubt“ (Splitvariante b) sein sollten.

Wichtig ist zu hinterfragen, ob der Grad an Routine nun Effekte, die sich aus der Formulierung der Antwortvorgaben ergeben, überdeckt, oder ob diese unabhängig davon auftreten. Um dazu eine Aussage liefern zu können, wurde eine logistische Regression berechnet (vgl. Tabelle 6).

Der Befund ist abermals durchaus bemerkenswert. So geht von der den Zielpersonen angebotenen Antwortvorgabe („verboten“ bzw. „nicht erlaubt“) tatsächlich ein deutlicher Effekt auf deren Antwortverhalten aus (vgl. die Variable *Split Formulierung* in Tabelle 6). Das Alter sowie die absolvierten Bildungsjahre besitzen hier einen Einfluss. Ohne klaren Effekt bleibt jedoch wiederum die Routine, die die Zielpersonen im Umgang mit ihrem Mobilfunkgerät haben.

Auch hier wurden wieder die bekannten Interaktionsterme in das Modell aufgenommen. Alle drei bleiben allerdings ohne Einfluss. Der Einfluss sowie die Erklärungskraft des Modells, wie bereits von allen vorangegangenen, sind äußerst gering.

Tabelle 6 Ergebnisse einer binär logistischen Regression zur Erklärung des Antwortverhaltens bei Unterschieden in der Frageformulierung, Effektkoeffizienten

Konstante	3,78**
Routine	1,15
Anwesenheit Dritter (Ref. Kat. = niemand anwesend)	0,98
Ort des Interviews (Ref. Kat. = zu Hause)	1,02
Alter	1,03***
Geschlecht (Ref. Kat. = männlich)	1,29
Bildung	0,94*
Routine x Anwesenheit Dritter	1,05
Routine x Ort des Interviews	1,15
Split Frageformulierung	1,79***
Routine x Split Frageformulierung	0,85

$n = 774$, $Pseudo-R^2 = 0,05$, * = 0,05, ** = 0,01, *** = 0,001;

Abhängige Variable: Frageformulierung, Ref. Kat. = Zustimmung

6 Zusammenfassung

Nachdem insgesamt sechs verschiedenen Kriterien zur Antwortqualität nachgegangen wurde, kann ein recht differenziertes Urteil zum Einfluss der Routine im Handygebrauch auf die Antwortqualität bei einer Befragung über das Mobilfunknetz gefällt werden. Dazu ist zu sagen:

Die Botschaft für die Mobilfunkbefragungen ist zunächst eine sehr positive. So dürfte die unterschiedliche Routine, die im Umgang mit dieser noch relativ neuen Technik in der Bevölkerung besteht, nicht zu deutlich verzerrten Umfrageergebnissen führen. Offenbar stellt die Mobilfunktechnik nicht so starke technische Ansprüche an den Benutzer, dass dadurch dessen Telefonverhalten bei Umfragen negativ beeinflusst würde. Dies steht im Kontrast zu über das Internet vermittelten Umfragestudien, wo solche Effekte aufgetreten sind. Das Internet ist (noch) schwieriger handhabbar als das Mobiltelefon. Für diesen Befund sprechen auch die insgesamt erfreulich geringen Werte für die erklärte Varianz (R^2 bzw. Nagelkerkes R^2). Dies bedeutet, dass vor allem modellexterne Variablen für das Antwortverhalten verantwortlich gemacht werden können.

Ein weiterer, in diesem Zusammenhang zukünftig stärker zu beachtender Aspekt dürfte die Qualität des Fragebogens sein. Der Erhebungsstandard der CELLA 1 Untersuchung enthält durchschnittlich komplizierte Fragestellungen. Der Mittelwert bei einer Frage zur Einschätzung der von den Zielpersonen empfundenen Schwierigkeit (v28) liegt auf einer fünfstufigen Skala bei 1,6 (alle 2.171 Befragten). Das heißt, die Erhebung wird zwischen „überhaupt nicht“ (Skalenwert 1) und „kaum schwierig“ (Skalenwert 2) eingeschätzt. Wie sich die Situation dagegen bei schwieriger zu beantwortenden Fragen oder bei längeren Umfragen verhält, muss an dieser Stelle offen bleiben.

Alle telefonischen Befragungen – zumal solche über den Mobilfunk – stellen einen Eingriff in die Privatsphäre dar. Nicht ausgeschlossen werden konnte deshalb, dass es gerade im Mobilfunkmodus, beispielsweise aufgrund von diversen Umgebungseinflüssen, zu einer starken Ablenkung der Zielpersonen kommen kann. Dies könnte zu stärkeren kognitiven Anstrengungen bei einer Befragung führen und sich damit negativ auf die Qualität des Antwortverhaltens auswirken. Die Befunde stützen eine solche Vermutung jedoch nicht. Bedenkt man, dass die Zielpersonen bei der Befragung über den Mobilfunkmodus zumeist auch zu Hause (vgl. die Ergebnisse von Frage v21 sowie Häder et al. 2009: 77ff.) erreicht worden sind, so ist dieser Befund durchaus sehr plausibel.

Die für die Suche nach möglichen Zusammenhängen zwischen der Routine und dem Antwortverhalten bei CELLA 1 eingesetzten vielfältigen methodischen Instrumente haben sich als ausreichend sensibel erwiesen. So zeigen insbesondere die Splitvarianten zu den Fragereihenfolgeeffekten sowie zum Effekt der Frageformulierung genau die erwarteten Muster. Auch die Fragestellungen zur Erinnerungsleistung und zum Pseudo-Opinion haben sich als besonders aussagestarke Indikatoren für die Antwortqualität erwiesen (vgl. die entsprechenden Tabellen). Der in der Regel nicht erfolgte Nachweis eines Effekts der Routine auf die Qualität des Antwortverhaltens deutet damit höchst wahrscheinlich nicht auf einen methodischen Fehler bzw. auf einen Zufall hin, sondern spiegelt die Realität richtig wider.

Ein weiteres Argument zur Erklärung der fehlenden Effekte könnte sein, dass in die Stichprobe vor allem Personen mit einer relativ hohen Routine gelangt sind, beispielsweise, weil andere ihre Teilnahme verweigert haben. Dass dies jedoch nicht so ist, konnte in Abbildung 1 gezeigt werden.

Welche Schlussfolgerungen können nun aus den Befunden von CELLA 1 für die Diskussion um die Qualität der Antwortfindung bei Umfragen allgemein gezogen werden? Erstens soll nochmals unterstrichen werden, dass bei nahezu allen Tests keine Effekte der Routine auf das Antwortverhalten ermittelt werden konnten.

Für eine ganze Reihe Annahmen gab es keine empirische Unterstützung anhand der CELLA 1 Daten. Damit muss davon ausgegangen werden, dass die Routine im Umgang mit dem Kommunikationskanal, über den die Befragung geführt wird, insgesamt nur einen relativ schwachen Einfluss auf die Antwortqualität hat. So spielen zum Beispiel der Bildungsstand der Zielperson, deren Alter, die Reihenfolge, in der die Fragen gestellt werden, gegebenenfalls das Geschlecht der Zielperson sowie die konkrete Formulierung der Antwortvorgaben eine stärkere Rolle als die Routine im Umgang mit dem Handy. Dies könnte Anlass dazu geben, insgesamt die Rolle der Routine im Umgang mit dem Vermittlungskanal einer Befragung – hier mit der Mobilfunktechnik – unter den Determinanten der Antwortqualität zu relativieren.

Die empirischen Ergebnisse geben Anlass zu der Vermutung, dass die Routine im Umgang mit dem Mobiltelefon nur dann Bedeutung für die Qualität des Antwortverhaltens erlangt, wenn technische Probleme mit dem Gerät auftreten. Da dies in der Regel offenbar nicht der Fall ist, laufen bei einer Mobilfunkbefragung ähnliche Prozesse ab, wie bei einer Erhebung über das Festnetz. Weder die Motivation, sich kognitiv mit der jeweiligen Fragestellung auseinander zu setzen, noch die Ablenkung infolge von Umgebungseinflüssen während der Befragung unterscheiden sich in Abhängigkeit von der Routine.

Jedoch sind Replikationsstudien erforderlich, um beispielsweise auch den Einfluss verschiedener Befragungsthemen als Bestimmungsgrößen für die Antwortqualität zu ermitteln. So ist nicht auszuschließen, dass bei kognitiv anspruchsvolleren Fragestellungen auch die Routine als Bestimmungsgröße für das Antwortverhalten wieder etwas an Einfluss gewinnt.

Literatur

- Alvarez, R. M. und J. Brehm, 1995: American Ambivalence Towards Abortion Policy: Development of a Heteroskedastic Probit Model of Competing Values. *American Journal of Political Science* 39: 1055-1082.
- Alvarez, R. M. und J. Brehm, 2002: *Hard Choices, Easy Answers: Values, Information, and American Public Opinion*. Princeton: Princeton University Press.
- Bishop, G. F., A. J. Tuchfarber und R. W. Oldendick, 1986: Opinions on Fictitious Issues: The Pressure to Answer Survey Questions. *Public Opinion Quarterly* 50: 240.
- Bishop, G. F., 1987: Experiments with the Middle Response Alternative in Survey Questions. *Public Opinion Quarterly* 51 (2): 220-232.
- Blumberg, S. J., und J. V. Luke, 2008: *Wireless Substitution: Early Release of Estimates from the National Health Interview Survey, January - June 2008*. Atlanta: CDC. <http://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless200812.pdf> (04.07.2010).
- Böcker, F., 1988: Scale Forms and their Impact on Rating's Reliability and Validity. *Journal of Business Research* 17 (1): 15-26.

- Christian, L. M., D. A. Dillman und J. D. Smyth, 2008: The Effects of Mode and Format on Answering to Scalar Questions in Telephone and Web Surveys. S. 250-274 in: J. M. Lepkowski, C. Tucker, J. M. Brick, E. de Leeuw, L. Japac, P. J. Lavrakas, M. W. Link und R. L. Sangster (Hg.): *Advances in Telephone Survey Methodology*. New York: Wiley.
- Converse, P. E., 1964: The Nature of Belief Systems. S. 206-261 in: D. E. Apter (Hg.): *Ideology and Discontent*. New York: Free Press.
- Converse, P. E., 1970: Attitudes and Non-attitudes: Continuation of a Dialogue. S. 186-189 in: E. R. Tuftte (Hg.): *The Quantitative Analysis of Social Problems*. Reading: Addison-Wesley.
- Craig, S. C., J. G. Kane und M. D. Martinez, 2000: Ambivalence, Attitude Strength, and Response Instability: A Two-Wave Panel Study of Abortion Attitudes in Florida. Washington, DC: American Political Science Association Meetings.
- Dillman, D. A., 2000: *Mail and Internet Surveys. The Tailored Design Method*. New York: Wiley.
- Dillman, D. A., T. L. Brown, J. Carlson, E.H. Carpenter, F. O. Lorenz, M. Robert, S. John und R. L. Sangster, 1995: Effects of Category Order on Answers to Mail and Telephone Surveys. *Rural Sociology* 60 (4): 674-687.
- Driscoll, D. L.; 2007: *Paraphrase: Write it in Your Own Words*. Purdue University Online Writing Lab.
- Faulbaum, F., P. Prüfer und M. Rexroth, 2009: *Was ist eine gute Frage? Die systematische Evaluation der Fragenqualität*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Gabler, S. und S. Häder, 2009: Gewichtung für die CELLA Studie. S. 51-55 in: M. Häder und S. Häder (Hg.): *Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Glendall, Ph. und J. Hoek, 1990: A Question of Wording. *Marketing Bulletin* 1: 25-36.
- Graeske, J. und T. Kunz, 2009: Stichprobenqualität der CELLA-Studie unter besonderer Berücksichtigung der Mobile-onlys. S. 57-70 in: M. Häder und S. Häder (Hg.): *Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Häder, M. und S. Häder (Hg.), 2009: *Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Häder, S., S. Gabler und C. Heckel, 2009: Stichprobenziehung für die CELLA Studie. S. 21-49 in: M. Häder und S. Häder (Hg.): *Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Häder, S., M. Häder, J. Graeske, T. Kunz und G. Schneiderat, 2009: Realisierung der Stichprobe. S. 71-82 in: M. Häder und S. Häder (Hg.): *Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Häder, M. und M. Kühne, 2009a: Theoretischer Rahmen und Untersuchungsdesign. S. 165-174 in: M. Häder und S. Häder (Hg.): *Telefonumfragen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Häder, M. und M. Kühne, 2009b: Erinnerungsleistung der Befragten in beiden Modes. S. 225-228 in: M. Häder und S. Häder (Hg.): *Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Häder, M. und M. Kühne, 2009c: Fragereihenfolgeeffekte. S. 207-216 in: M. Häder und S. Häder (Hg.): *Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Hippler, H. J. und N. Schwarz, 1986: Not Forbidding isn't Allowing: The Cognitive Basis of the Forbid – Allow Asymmetry. *Public Opinion Quarterly* 50 (1): 87-96.

- Holleman, B., 2000: The Forbid/Allow Asymmetrie. On the Cognitive Mechanism Underlying Wording Effects in Surveys. Amsterdam Atlanta: Radopi.
- Hyman, H. H. und P. B. Sheatsley, 1950: The Current Status of American Public Opinion. S. 11-34 in: Payne, J. C. (Hg.): The Teaching of Contemporary Affairs. Twenty-first Yearbook of the National Council of Social Studies. New York: National Education Association.
- Jobe, J. B. und D. J. Herrmann, 1996: Implications of Models of Survey Cognition for Memory Theory. S. 193-205 in: D. J. Herrmann, M. Johnson, Ch. Herzog und P. Hertel (Hg.): Basic and Applied Memory Research. Bd. 2. Practical Applications. Hillsdale, NJ: Erlbaum.
- Kriesi, H., 2004: Stability and Change of Opinion: The Case of Swiss Policy Against Pollution Caused by Cars. S. 242-267 in: W. E. Saris und P. M. Sniderman (Hg.); Studies in Public Opinion: Gauging Attitudes, Nonattitudes, Measurement Error and Change. Princeton: Princeton University Press.
- Krosnick, J. A. und D. F. Alwin, 1987: An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement. *Public Opinion Quarterly* 51 (2): 201-219.
- Krosnick, J. A., 2000: The Threat of Satisficing in Surveys: The Shortcuts Respondents Take in Answering Questions. *Survey Methods Newsletter* 20 (1): 4-8.
- Krosnick, J. A., A. L. Holbrook, M. K. Berent, R. T. Carson, W. M. Hanemann, R. J. Kopp, R. C. Mitchell, S. Presser, P. A. Ruud, V. K. Smith, W. R. Moody, M. C. Green und M. Conaway, 2002: The Impact of No Opinion Response Options on Data Quality: Non-Attitude Reduction Or an Invitation to Satisfice? *Public Opinion Quarterly* 66: 371-403.
- Kuusela, V., M. Callegaro und V. Vehovar, 2008: The Influence of Mobile Telephones on Telephone Surveys. S. 87-112 in: J. M. Lepkowski, C. Tucker, J. M. Brick, E. de Leeuw, L. Japac, P. J. Lavrakas, M. W. Link und L. Sangster (Hg.): *Advances in Telephone Survey Methodology*, New Jersey: Wiley.
- Kühne, M. und R. Böhme, 2006: Effekte von Informationsstand, Wissen und Einstellungsstärke von Befragten auf die Antwortstabilität. *ZUMA-Nachrichten* 59: 42-71. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/zuma_nachrichten/zn_59.pdf (22.07.2010).
- Kühne, M. und M. Häder, 2009: Konsistenzen und Inkonsistenzen im Antwortverhalten. S. 217-224 in: M. Häder und S. Häder (Hg.): *Telefonbefragungen über das Mobilfunknetz. Konzept, Design und Umsetzung einer Strategie zur Datenerhebung*. Wiesbaden: VS Verlag für Sozialwissenschaften.
- de Leeuw, E. D. und J. H. Jopp, 2008: The Cornerstones of Survey Research. S. 1-17 in: E. D. de Leeuw, J. H. Joop und D. A. Dillman (Hg.): *International Handbook of Survey Methodology*. New York: Lawrence Erlbaum Associates.
- Loosveldt, G., 1997: Interaction Characteristics in Some Question Wording Experiments. In: *Bulletin de Méthodologie Sociologique (BMS)* 56: 20-31.
- Mayerl, J. und D. Urban, 2008: Antwortreaktionszeiten in Survey-Analysen. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Payne, S. L., 1949: Case Study in Question Complexity. *Public Opinion Quarterly* 13 (4): 653-658.
- Reuband, K. H., 2000: „Pseudo Opinions“. *ZA-Information* 46: 26-38. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/za_information/ZA-Info-46.pdf (22.07.2010).
- Reuband, K. H., 2003: Variationen der Permissivität: Wie Frageformulierungen unterschiedliche Antwortverteilungen erbringen, wenn von 'Erlauben' oder 'Verbieten' die Rede ist. *ZA-Information* 53: 86-96. http://www.gesis.org/fileadmin/upload/forschung/publikationen/zeitschriften/za_information/ZA-Info-53.pdf (22.07.2010).
- Rugg, D. (1941): Experiments in Wording Questions. *Public Opinion Quarterly* 5 (1):91-92.
- Schuman, H., 1992: Context Effects: State of the Past / State of the Art. S. 5-20 in: N. Schwarz und S. Sudman (Hg.): *Context Effects in Social And Psychological Research*. New York: Springer.

- Schuman, H. und S. Presser, 1981: Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context. New York: Academic.
- Schwarz, N., H.-J. Hippler, B. Deutsch und F. Strack, 1985: Response Scales: Effects of Category Range on Reported Behaviour and Comparative Judgements. *Public Opinion Quarterly* 49 (3): 388-395.
- Shoemaker, P. J., M. Eichholz und E. A. Skewes, 2002: Item Nonresponse: Distinguishing Between Don't Know and Refuse. *International Journal of Public Opinion Research* 14 (4): 193-201.
- Slaby, M., 1998: Eine Untersuchung zur Konstanz des Meinungsurteils von Befragten im Interviewverlauf dargestellt am Fallbeispiel „Bewertung der Gentechnik“. Stuttgart: IfS.
- Sudman, S., N. M. Bradburn und N. Schwarz, 1996: Thinking about Answers. The Application of Cognitive Processes to Survey Methodology. San Francisco: Jossey-Bass.
- Stadler, K., 1983: Die Skalierung in der empirischen Forschung. München: Infratest.
- Strack, F., 1994: Zur Psychologie der standardisierten Befragung. Kognitive und kommunikative Prozesse. Berlin: Springer.
- Strack, F., L. L. Martin, 1988: Thinking, Judging, and Communicating: A Process Account of Context Effects in Attitude Surveys. S. 123-148 in: H.-J. Hippler (Hg.): *Social Information Processing and Survey Methodology*. 2. Auflage. New York: Springer.
- Tourangeau, R., L. J. Rips und K. A. Rasinski, 2000: *The Psychology of Survey Response*. Cambridge: University Press.
- Waterplas, L., J. Billet und G. Loosveldt, 1988: De verbieden versus niet toelaten asymmetrie. Een stabiele formuleringseffect in survey-onderzoek? *Mens en Maatschappij* 63: 399-417.
- Zaller, J. R., 1992: *The Nature and Origins of Mass Opinion*. Cambridge: Cambridge UP.

Anschrift der Autoren

Prof. Dr. Michael Häder
Technische Universität Dresden
Philosophische Fakultät, Institut für Soziologie
Lehrstuhl für Methoden der empirischen Sozialforschung
01062 Dresden
michael.haeder@tu-dresden.de

Dr. Mike Kühne
Technische Universität Dresden
Philosophische Fakultät, Institut für Soziologie
Lehrstuhl für Methoden der empirischen Sozialforschung
01062 Dresden
mike.kuehne@tu-dresden.de