

Heikle Fragen in mündlichen Interviews:
Ergebnisse einer Methodenstudie im studentischen Milieu *

Peter Preisendörfer
Universität Mainz
Institut für Soziologie

Mainz, Juli 2008

* Das Arbeitspapier entstand im Rahmen des Forschungsprojekts „Asking Sensitive Questions“ (Teilprojekt PR 237/6-1), das ein Verbundprojekt der ETH-Zürich, der Universität Leipzig und der Universität Mainz ist und von der Deutschen Forschungsgemeinschaft (DFG) im Rahmen des Schwerpunktprogramms „Survey Methodology“ gefördert wird.

Heikle Fragen in mündlichen Interviews:
Ergebnisse einer Methodenstudie im studentischen Milieu

Zusammenfassung: In einem „sensitive topic survey“ mit Studierenden der Universität Mainz (n=578) wurde die Brauchbarkeit von drei Techniken zur Erhebung heikler Sachverhalte untersucht: Wording/Framing-Techniken, die Technik des vertraulichen Kuverts und die Randomized-Response-Technik. Bezüglich Wording/Framing bestätigt sich der Verdacht, dass diese in der Lehrbuchliteratur viel zitierten Techniken oft nicht halten, was sie versprechen. Demgegenüber erweist sich die Technik des vertraulichen Kuverts im Anwendungsfall der Erhebung sexuellen Verhaltens als hilfreich. Die eingesetzte Randomized-Response-Technik in der Variante von „forced response“ brachte nicht die erwarteten Ergebnisse. Der Beitrag ist insgesamt ein erster Schritt im Rahmen eines breiteren Forschungsprogramms zur Thematik heikler Fragen in verschiedenen Arten von Befragungen.

Sensitive Questions in Face-to-Face Interviews:
Findings of a Methodological Study with University Students

Abstract: A “sensitive topic survey” among students of the University of Mainz (n=578) was conducted to evaluate the merits of three techniques to investigate sensitive behavior: wording/framing techniques, the sealed envelope technique, and the randomized response technique. Concerning wording/framing, the results show that these techniques, which can be found in all textbooks of social research methods, often do not meet their expectations. The sealed envelope technique proved to be useful in asking questions about sexual behavior. Not at all successful was the randomized response technique in the form of “forced response”, developed and tested in the student survey. The article is a first contribution to a broader research programme about sensitive topics in different survey modes.

Heikle Fragen in mündlichen Interviews: Ergebnisse einer Methodenstudie im studentischen Milieu

1. Einführung

Heikle Fragen sind eine besondere Herausforderung für die Theorie und Praxis der Umfrageforschung. Denn viele Probleme, die generell in Befragungen auftreten (wie Antwortverweigerungen oder die Tendenz zu sozial erwünschten Antworten), stellen sich bei „sensitive topics“ in verschärfter Form. Dies hat es mit sich gebracht, dass gerade die Erhebung heikler Sachverhalte empirische Sozialforscher/innen dazu angespornt hat, Methodenexperimente zu starten, unterschiedliche Befragungsformate zu vergleichen und methodische Innovationen zu erproben. Von diesen methodischen Innovationen profitiert dann im Endergebnis auch die konventionelle Survey-Methodologie.

In der Tat gibt es zur Thematik heikler Fragen eine beachtliche internationale Forschung. Beginnend mit dem satirischen Kurzbeitrag „Asking the Embarrassing Question“ von Allen Barton (1958) hat sich im Verlauf der Jahre ein eigenständiges Forschungsfeld zu heiklen Fragen etabliert: mit einschlägigen Sammelbänden (z.B. Renzetti/Lee 1993), Lehrbüchern (z.B. Fox/Tracy 1986; Lee 1993), Überblicksartikeln (z.B. Barnett 1998; Tourangeau et al. 2000: Kap. 9; Tourangeau/Yan 2007) und auch nennenswerten Beiträgen aus Disziplinen außerhalb der Sozialwissenschaften (z.B. aus der medizinischen Forschung). Die zwei wichtigsten Arbeitsbereiche im Forschungsfeld um heikle Fragen sind zum einen konzeptionell-theoretische Problemstellungen, zum anderen konkrete Methoden und Techniken, um heikle Sachverhalte zu erheben. Auf der konzeptionell-theoretischen Ebene wird diskutiert, was überhaupt heikle Fragen sind, auf der Grundlage welcher Eigenschaften Fragen einen heiklen Charakter gewinnen, wie sich die Vorstellungen von „heikel“ im Zeitablauf verändert haben, inwieweit es im Vergleich verschiedener Gesellschaften Unterschiede gibt, und bei welchen Fragen in welchem Ausmaß mit Antwortverzerrungen (typischerweise in der Form von „underreporting“) zu rechnen ist. Mit Blick auf Techniken und Erhebungsmethoden geht es darum, Wege zu finden, mit denen sich heikle Sachverhalte valider als durch „direktes Abfragen“ erfassen lassen. Dabei wird ein Arsenal an Methoden diskutiert, das von einfachen Formulierungstricks, über vertrauliche Kuverts und Vignetten, bis hin zur sogenannten Randomized-Response-Technik reicht (für einen knappen und sehr schönen Überblick vgl. Lee 1993: Kap. 5). Teile dieses Methodenarsenals werden standardmäßig auch in den gängigen Lehrbüchern zur empirischen Sozialforschung angesprochen (z.B. Schnell 2005: 339ff.; Diekmann 2007: 447ff. und 488ff.).

Enttäuschend schwach und spärlich ist allerdings die bisherige Primärforschung zu heiklen Fragen im deutschsprachigen Raum. Ein früher Appell von van Koolwijk (1969), der ein diesbezügliches Forschungsprogramm anmahnte, ist weitgehend ungehört verhallt. Über die Zeit hinweg finden sich nur wenige und sporadische Einzelbeiträge (u.a. Deffaa 1982; Becker/Günther 2004). Diese schwache Profilierung der deutschsprachigen Forschung war für uns Anlass, im Rahmen des DFG-Schwerpunktprogramms „Survey Methodology“ ein Teilprojekt zu initiieren, das sich – ab Anfang 2008 – über mehrere Jahre hinweg mit der Theorie und Empirie heikler Fragen befassen soll.¹

Der vorliegende Beitrag ist ein erstes Ergebnis der Arbeiten in diesem Projekt. Es sollen die Befunde einer mündlichen Befragung von Studierenden der Universität Mainz vorgestellt werden, in der drei in der einschlägigen Literatur viel zitierte Techniken zur Erhebung heikler Sachverhalte erprobt wurden: 1) Wording/Framing-Techniken, 2) die Technik des vertraulichen Kuverts und 3) die Randomized-Response-Technik. Im folgenden Abschnitt (Abschnitt 2) wird zunächst die empirische Datenbasis kurz erläutert. Anschließend werden die mit den drei Techniken erzielten Ergebnisse dargestellt. Die heiklen Themenbereiche, auf welche die Fragen abstellten, unterschieden sich bei den drei Techniken: Bei den Wording/Framing-Techniken waren Alkohol und Drogen sowie kleinere kriminelle Alltagsdelikte der inhaltliche Gegenstand (Abschnitt 3). Bei der Technik des vertraulichen Kuverts ging es um sexuelle Erfahrungen und sexuelles Verhalten (Abschnitt 4). Und bei der Randomized-Response-Technik wurde nach unlauteren Praktiken beim Studium gefragt (Abschnitt 5). Die mit den genannten Techniken erzielten Ergebnisse werden stets mit Ergebnissen verglichen, die bei einer Teilstichprobe mittels direkten Fragens erzielt wurden. Im Schlussabschnitt (Abschnitt 6) wird versucht, die „Lehren“ aus dem Projekt in wenigen Worten zusammenzufassen, zudem werden Hinweise auf weiteren Forschungsbedarf gegeben.

¹ Die Forschungsprojektgruppe besteht aus einem Team, das an den Soziologischen Instituten der Universität Leipzig (Ivar Krumpal, Karl-Dieter Opp, Thomas Voss), der Universität Mainz (Peter Preisendörfer, Felix Wolter) und der ETH-Zürich (Elisabeth Coutts, Andreas Diekmann) angesiedelt ist.

2. Empirische Datenbasis

Die Daten, die hier verwendet werden, stammen aus einem Lehr-Forschungsprojekt, das vom Autor in der Form einer zweisemestrigen Veranstaltung am Institut für Soziologie der Universität Mainz initiiert und geleitet wurde. Im Rahmen dieses Projekts wurden in den Monaten Oktober und November 2007 insgesamt 578 Studierende der Universität Mainz in halbstündigen face-to-face Interviews befragt. Die Befragung basierte auf einer einfachen Zufallsstichprobe, bei der – mit den zwei Beschränkungen, dass es sich erstens um reguläre Studierende (also ohne Austauschstudenten, Promovenden, Gasthörer u.ä.) und zweitens um Studierende mit Hauptwohnsitz in der Stadt Mainz handeln musste – ursprünglich 1508 Adressen aus der Datei des Studierendensekretariats gezogen wurden.² Jede/r der 47 studentischen Projektteilnehmer/innen erhielt 30 Adressen und sollte versuchen, damit 15 Interviews zu realisieren. Der Arbeitseinsatz und der Erfolg der studentischen Interviewer/innen war unterschiedlich, wobei einige nicht alle ihre Adressen ausschöpften, andere zusätzliche Adressen brauchten und von den meisten die Zielvorgabe von 15 Interviews nicht ganz erreicht wurde. Lässt man die Adressen außer Acht, die entweder überhaupt nicht bearbeitet wurden oder definitiv den Charakter von stichprobenneutralen Ausfällen hatten (z.B. weil die Personen ihr Studium inzwischen abgebrochen hatten), ergibt sich eine Ausschöpfung von 43%, was man sicher nicht als ein gutes Ergebnis werten kann.

Die erste Kontaktaufnahme mit den Zielpersonen erfolgte durch ein Anschreiben. In diesem Anschreiben wurde die Studie unter dem Titel „Campus and Beyond“ vorgestellt – mit der Erläuterung, dass es in der Befragung um studentisches Leben an der Universität und in der Stadt Mainz ginge, und mit dem expliziten Vermerk, dass dabei auch Themen angesprochen würden, „über die man nicht so oft und gerne redet“. Tatsächlich hatte die Befragung den Charakter eines sogenannten „sensitive topic survey“, d.h. es ging hauptsächlich um heikle Themen. Es ist nicht auszuschließen, dass dieser Hinweis auf heikle Themen bei einigen Studenten bewirkt hat, dass sie eine Teilnahme an der Befragung verweigerten, sodass es zu einer Selbst-Selektion in der Stichprobe kam. Da der Hinweis freilich behutsam nur in einem Nebensatz erfolgte und Studenten insgesamt eine Bevölkerungsgruppen sind, die überdurchschnittlich offen und aufgeschlossen sind, kann man annehmen, dass die vermutete

² Eine weitere Beschränkung der Stichprobe war, dass die Zielpersonen (laut Information in der Datei des Studierendensekretariats) in den Monaten Mai bis Oktober geboren sein mussten. Die Interviewer/innen wussten von dieser Beschränkung nichts und erfragten in ihren Interviews auch den Geburtsmonat und das Geburtsjahr der Zielperson. Auf diese Weise sollten Hinweise auf gefälschte Interviews gewonnen werden. Tatsächlich lieferten drei Interviewer/innen jeweils ein Geburtsdatum außerhalb der Zeitspanne Mai bis Oktober. Die Feldarbeit dieser drei Interviewer/innen wurde genauer überprüft, ohne dass sich nennenswerte Verdachtsmomente ergaben.

Selbst-Selektion kaum gravierend ist. Im Anschreiben wurde angekündigt, in den nächsten Tagen werde ein/e (namentlich genannte/r) studentische/r Interviewer/in vorbeikommen, was dann in der Regel auch geschah.

Den Fragebogen gab es in zwei Varianten: In Variante 1, der „direkten Version“, wurden alle interessierenden heiklen Sachverhalte direkt bzw. möglichst direkt abgefragt. In Variante 2, der „indirekten Version“, kamen die oben schon genannten Techniken zum Einsatz, also die Wording/Framing-Techniken (WFT), die Technik des vertraulichen Kuverts (VKT) und die Randomized-Response-Technik (RRT). Nach dem Muster von Split-Ballot sollte die direkte Version die Baseline stellen, gegen die die Ergebnisse mit der indirekten Version zu vergleichen waren. Die Erwartung war, dass die indirekte Version in der Regel höhere Prozentwerte für das heikle Verhalten liefert als die direkte Version. Ausgehend von der Zielgröße, dass rund ein Drittel der realisierten Interviews direkt und zwei Drittel indirekt sein sollten, wurde mit einem Zufallsverfahren (Randomisierung) ein Drittel der studentischen Interviewer/innen dazu bestimmt, direkte Interviews zu führen; die verbleibenden zwei Drittel mussten indirekte Interviews führen. Die durchgängige Verpflichtung der Interviewer entweder auf direkte oder aber auf indirekte Interviews birgt zwar die Gefahr einer Konfundierung von Interviewermerkmalen und Befragungstechnik in sich, aber durch die Zufallszuweisung dürfte diese Gefahr weitgehend begrenzt worden sein. Im Endergebnis waren von den 578 Interviews, die geführt wurden, 211 (37%) direkt und 367 (63%) indirekt. Signifikante Unterschiede in der Erfolgs- bzw. Rücklaufquote bei den direkten und indirekten Interviews gab es nicht.

Trotz der heiklen Inhalte der Erhebung kamen Interviewabbrüche so gut wie nicht vor (vier Fälle). Wie auch in anderen Umfragen waren zum einen die Nichterreichbarkeit der Zielpersonen und zum anderen deren anfängliche Entscheidung, teilzunehmen oder nicht, die entscheidenden Barrieren. Die Interviewer/innen stuften die Bereitschaft der 578 erfolgreich Befragten, die Fragen zu beantworten, zu 90% als gut, zu 9% als mittelmäßig und zu 1% als schlecht ein. Weiterhin werteten sie auf einer fünfstufigen Skala (überhaupt nicht zuverlässig bis sehr zuverlässig) die Angaben der Befragten zu 89% als zuverlässig bzw. sehr zuverlässig.

3. Ergebnisse mit den Wording/Framing-Techniken (WFT)

Die elementarste und häufigste Technik, die für eine im Vergleich zur direkten Methode validere Erhebung heikler Sachverhalte vorgeschlagen wird, lässt sich unter dem Stichwort „Wording/Framing“ zusammenfassen. Tatsächlich handelt es sich bei WFT um eine ganze Gruppe von Techniken. Dabei wird mit verharmlosenden Formulierungen gearbeitet, mit Suggestivfragen, Unterstellungen, mit dem sogenannten Mitläufereffekt, mit zusätzlichen Informationen, die das angezielte Verhalten als normal oder selbstverständlich erscheinen lassen sollen, und mit ähnlichen „Tricks“ einer mehr oder weniger raffinierten Wortwahl sowie einer geschickten Einbettung und Kontextualisierung der kritischen Fragen (als Klassiker dazu vgl. vor allem Barton 1958; Sudman/Bradburn 1974, 1982; Bradburn/Sudman 1979).

Wie schon eingangs erwähnt, bezogen sich die Fragen, mit denen wir in unserer Studie den „Ertrag“ von WFT ausloten wollten, inhaltlich zum einen auf Alkohol/Drogen, zum anderen auf kleinere kriminelle Alltagsdelikte. Bei Alkohol/Drogen wurden beschwichtigende Zusatzinformationen geliefert, und weiterhin wurde der Mitläufereffekt zu aktivieren versucht. Bei den kriminellen Alltagsdelikten wurde mit der sogenannten Trichter-Technik gearbeitet, d.h. in den indirekten Interviews wurde ausgehend von allgemeinen Fragen langsam und behutsam zu den eigentlich interessierenden heiklen Fragen hingeführt, was als eine spezielle Form von Einbettung bzw. Framing gesehen werden kann. Insgesamt enthielt die Umfrage in diesem Bereich sieben Kernfragen.³ Die Auszählungen dieser Fragen im Vergleich von direkter und indirekter Fragebogenversion sind in Tabelle 1 zusammengestellt.

Tabelle 1 hier

Aus der Tabelle und ihrem erläuternden „Unterbau“ lässt sich Folgendes ablesen: Einen häufigen Alkoholkonsum (mehr als dreimal pro Woche) berichten nur wenige Student/innen. Der Hinweis auf die angeblich positiven Effekte mäßigen Alkoholkonsums bei der WFT-Frage trägt nicht dazu bei, dass das Trinken von Alkohol häufiger angegeben wird. Dasselbe gilt für das Eingeständnis von Vollrausch-Situationen, bei denen man/frau „so richtig betrunken“ ist. Die direkten Prozentwerte und die bei WFT, wo das Trinken über den Durst als besonderes „Event“ verpackt wurde, unterscheiden sich nicht. Eigentlich bemerkenswert

³ Zusätzlich zu den Kernfragen gab es diverse weitere Fragen, die auf inhaltliche Details abzielten oder aber nur Teilgruppen der Befragten ansprachen. Die Ergebnisse dieser Zusatzfragen werden hier nicht berichtet, zumal sie das Gesamtbild der Befunde nicht verändern. Diese Beschränkung in der Darstellung der Ergebnisse gilt auch für die späteren Abschnitte 4 und 5.

ist hier nur, wie häufig das Vollrausch-Trinken im studentischen Milieu offenbar vorkommt. Keine signifikante Differenz zwischen direkt und WFT ergibt sich auch bei der Frage, ob die Zielpersonen jemals illegale Drogen genommen haben. WFT mit dem Vorspann, dass auch viele VIPs in ihrer Jugend illegale Drogen probiert haben, liefert zwar einen um fünf Prozentpunkte höheren Wert, aber ein Unterschied in dieser Größenordnung liegt im Bereich von Zufallsschwankungen.⁴

Bei den Fragen „GEZ-Betrug“ bis „strafrechtliches Verfahren“ wurde nach der Trichter-Technik vorgegangen. Die Strategie, mit allgemeinen Fragen zu beginnen und erst langsam „auf den Punkt zu kommen“, erweist sich lediglich im Fall des GEZ-Betrugs als erfolgreich. Hier gelingt es, bei denjenigen, die zur Entrichtung von Radio- und Fernsehgebühren verpflichtet sind (82% aller Befragten⁵), das Eingeständnis aktuell nicht-gezahlter Gebühren mittels Trichter im Vergleich zum direkten Vorgehen um neun Prozentpunkte zu steigern. Dass die Frage nach den GEZ-Gebühren tatsächlich heikel ist, heikel im Sinne einer Befürchtung unangenehmer persönlicher Konsequenzen, zeigte sich in unseren Interviews dergestalt, dass an dieser Stelle den Interviewern gegenüber wiederholt Skepsis geäußert wurde und es in einem Fall genau hier sogar zu einem Abbruch des Interviews kam. Keine bzw. nur marginale Unterschiede zwischen direkt und WFT ergeben sich hingegen bei den drei Fragen, ob man schon einmal ein Buß- oder Verwarnungsgeld für eine Verkehrsordnungswidrigkeit zahlen musste, ob man jemals einen Ladendiebstahl begangen hat und ob man schon irgendwann einmal in einem strafrechtlichen Verfahren angeklagt war.⁶

Insgesamt legen die Ergebnisse zu Wording und Framing die Schlussfolgerung nahe, dass man davon in der Regel nicht allzu viel erwarten kann und sollte. Diese Interpretation wird unterstützt, wenn man zwei Zusatzfragen auszählt, die am Ende unseres Interviews gestellt wurden: erstens in welchem Ausmaß empfanden die Zielpersonen die Fragen in den Bereichen Alkohol/Drogen und Alltagskriminalität als unangenehm (Antwortskala von 1=überhaupt nicht unangenehm bis 10=sehr unangenehm), zweitens in welchem Ausmaß haben sie nach eigenen Angaben die Fragen in den beiden Bereichen ehrlich beantwortet (Antwortskala von 1=überhaupt nicht ehrlich bis 10=sehr ehrlich). In Abhängigkeit von der

⁴ Bei den Signifikanztests in Tabelle 1 handelt es sich um einseitige Chi-Quadrat-Tests.

⁵ Studierende mit geringem „Einkommen“ können sich unter bestimmten Bedingungen von der Gebührenpflicht befreien lassen. Weiterhin leben einige der Befragten angeblich tatsächlich ohne Radio und Fernsehen. Die 18% ohne GEZ-Pflicht wurden aus der Analyse ausgeklammert.

⁶ Speziell bei der Ladendiebstahl-Frage wurde auch deutlich, dass bei einer Trichteranordnung der Fragen Vorsicht geboten ist: Wer in unseren Interviews am Anfang des Trichters die Auskunft gab, ein Ladendiebstahl sei unter keinen Umständen vertretbar (auch nicht im Fall einer materiellen Notlage), tut sich gewiss schwer damit anschließend zu bekennen, eventuell selbst schon einmal einen Ladendiebstahl begangen zu haben.

Befragungsvariante (direkt versus WFT) ergeben sich in den beiden Bereichen (Alkohol/Drogen und Alltagskriminalität) und auf den beiden Dimensionen (Unangenehmheit und Ehrlichkeit) keine signifikanten Differenzen. Umgekehrt gilt immerhin, dass Wording und Framing wohl auch nicht schaden, zumal die Richtung der Antwortveränderungen in Tabelle 1 konsistent den Erwartungen entspricht. Diese Techniken sind zwar anschaulich und prima facie einleuchtend (und eignen sich deshalb gut für die Lehrbuchliteratur), aber sie halten wohl nur in Ausnahmefällen, was sie versprechen. Dies möglicherweise auch deshalb, weil sich die Bevölkerung im Zeitalter der Markt- und Meinungsforschung inzwischen auf derartige Praktiken eingestellt hat.

4. Ergebnisse mit der Technik des vertraulichen Kuverts (VKT)

Ebenfalls noch eher zum Kanon konventioneller Techniken der Erhebung heikler Sachverhalte gehört die „sealed envelope technique“, auf deutsch die Technik des vertraulichen Kuverts (VKT). Sie findet sich als „sealed ballot“ schon bei Barton (1958) und kam beispielsweise auch im ALLBUS 2000 im Rahmen einer Erhebung delinquenter Verhaltensweisen zum Einsatz (Becker/Günther 2004; vgl. weiterhin Sudman/Bradburn 1974, 1982; De Leeuw 2001). Bei der VKT wird den Befragten an einer bestimmten Stelle des Interviews ein separater Kurzfragebogen übergeben, den sie eigenständig ausfüllen und dann in ein Kuvert eintüten sollen, sodass der Interviewer nichts über die Antworten weiß und mithin in diesem Bereich Anonymität hergestellt wird.

Die Interviewer/innen in unserer Erhebung hatten ein DIN-A4-Kuvert dabei, auf das groß und fett „Vertraulich“ gedruckt war. In dem Kuvert steckte ein leerer vierseitiger Fragebogen mit 20 Fragen, die sich alle mehr oder weniger direkt auf das Thema Sexualität bezogen. Die Interviewer/innen gaben die Anweisung: „Ich möchte dich bitten, die Fragen aus dem Kuvert herauszunehmen, sie eigenständig auszufüllen und dann wieder in das Kuvert zurückzugeben. Das Kuvert klebst du dann zu. Und nicht ich, sondern nur eine Person am soziologischen Institut, die dich nicht kennt, darf es öffnen. Dies gewährleistet die Anonymität deiner Antworten“. Die Auszählung der sechs Kernfragen zu sexuellen Erfahrungen und Verhaltensweisen der Respondenten findet sich in Tabelle 2. Wie schon in Tabelle 1 interessiert vor allem der Vergleich zwischen direkter und indirekter Fragebogenvariante (direkt versus VKT).

Tabelle 2 hier

Bei den ersten fünf Fragen in Tabelle 2 zeigen sich – wider Erwarten – keine signifikanten

Unterschiede zwischen den Abfragemodi „direkt versus VKT“.⁷ Im Schnitt hatten die Student/innen in ihrem bisherigen Leben vier (Median) bzw. sechs (arithmetisches Mittel) unterschiedliche Sexualpartner. Der etwas höhere Mittelwert bei VKT kommt hauptsächlich durch zwei „Ausreißer“ zustande, die beide eine Zahl von 100 Sexualpartnern angegeben haben. Die Häufigkeit des Geschlechtsverkehrs in den letzten vier Wochen liegt bei direkter und VKT-Abfrage übereinstimmend bei drei (Median) bzw. fünf Mal (arithmetisches Mittel). Dieser im Vergleich zu anderen Studien (z.B. Jasso 1985) sehr niedrige Wert resultiert in erster Linie daraus, dass viele der befragten Studierenden (42%) aktuell nicht in einer festen Partnerschaftsbeziehung leben. Bei den zwei Fragen, ob die Zielpersonen jemals einen sogenannten One-Night-Stand hatten und ihre aktuelle oder letzte feste Beziehung jemals mit einem sogenannten Seitensprung belastet haben, liefert VKT sogar etwas niedrigere Prozentwerte. Der Tendenz nach erwartungskonform ist das Antwortmuster wieder bei der Frage nach homosexuellen Kontakten, wenngleich weit entfernt von der 5%-Niveau-Signifikanzschwelle. Doch zu guter Letzt: Bei der Frage, die in „Sensitivity-Rankings“ in der Regel an oberster Stelle steht (vgl. z.B. bereits Bradburn/Sudman 1979: 17), nämlich der Masturbation, erweist sich VKT als eindeutig überlegen. Im direkten Fragemodus gestehen nur 15% ein, sich häufig bzw. sehr häufig selbst zu befriedigen, mittels VKT steigt der Wert auf 29%. Dieser Befund spricht für die Hypothese, dass speziell und vor allem bei hochsensiblen Themen VKT validere Antworten bringt. In der Tat lässt sich argumentieren, dass z.B. homosexuelle Kontakte oder auch One-Night-Stands bei Student/innen überhaupt nicht (mehr) den Charakter heikler Fragen haben. Und bei der Zahl der bisherigen Sexualpartner könnte der Wert 0 ähnlich heikel sein wie der Wert 100, d.h. die sozial erwünschte Antwort dürfte eher im mittleren Wertebereich liegen.

Obwohl VKT in unserer Erhebung lediglich bei der Selbstbefriedigung besser abschneidet, gibt es weitere gute Gründe für den Einsatz dieser Technik bei sensiblen Themen. Zweifellos ist es für die Interviewer/innen einfacher und problemloser, Sachverhalte wie die in Tabelle 2 mittels VKT abzufragen. Aus der einschlägigen Methodenforschung ist bekannt, dass heikle Fragen oft mehr ein Problem der Interviewer als der Befragten sind, d.h. die Interviewer haben Vorbehalte und Bedenken, diese Fragen zu stellen, weil sie negative Reaktionen seitens der Befragten befürchten (vgl. dazu Lee 1993: 101; Schnell 1997: 274f.; Hox/de Leeuw 2002; Schnell/Kreuter 2005). Wenn dem so ist, muss man damit rechnen, dass einige Interviewer die heiklen Fragen im direkten Fragemodus „überspringen“ und im Nachhinein selbstständig angebliche Antworten eintragen.

⁷ Bei der Zahl der Sexualpartner und der Häufigkeit des Geschlechtsverkehrs wurde die Signifikanz des Unterschiedes mittels t-Tests überprüft. Ansonsten handelt es sich bei den Signifikanztests (wie in Tabelle 1) um einseitige Chi-Quadrat-Tests.

Weiterhin kann man erwarten, dass die Quote der Antwortverweigerungen (Item-Nonresponse) bei VKT niedriger ausfällt als bei direkten Fragen. Bei den drei Items „One-Night-Stand“, „Seitensprung“ und „homosexuelle Kontakte“ lag der Anteil der Missings in unserer Erhebung unabhängig von der Frageform stets unter 2%, sodass sich signifikante Unterschiede im direkt-VKT-Vergleich von vorneherein nicht ergeben konnten. Die direkte Frage nach der Zahl der Sexualpartner lieferte 6,6% Missings, VKT hingegen lediglich 1,6%. Die entsprechenden Vergleichswerte bei der Häufigkeit des Geschlechtsverkehrs in den letzten vier Wochen sind 4,7% versus 1,9%, und bei der Selbstbefriedigung 5,7% versus 2,7%.

Auch beim Thema Sex wurde gegen Ende des Interviews gefragt, inwieweit die Fragen in diesem Bereich von den Respondenten erstens als unangenehm empfunden und zweitens ehrlich beantwortet wurden. Von den direkt Befragten stuften 20% die Sexfragen als „überhaupt nicht unangenehm“ ein (Wert 1 auf einer Skala von 1-10), von den VKT-Befragten hingegen 40%. Keine nennenswerte Differenz ergibt sich allerdings bei der deklarierten Ehrlichkeit der Antworten: Im direkten Fragemodus haben 79%, im VKT-Modus 81% angeblich „sehr ehrlich“ (Wert 10 auf einer Skala von 1-10) geantwortet.

5. Ergebnisse mit der Randomized-Response-Technik (RRT)

Im Rahmen des Methodenspektrums zur Erhebung heikler Sachverhalte haben insbesondere statistisch versierte Sozialforscher/innen vielfach eine Vorliebe für die Randomized-Response-Technik (RRT). Ursprünglich eingeführt von Stanley Warner (1965), wird seit nunmehr über 40 Jahren in empirischen Studien sporadisch mit RRT experimentiert. Trotzdem ist nach wie vor umstritten, ob und inwieweit die Methode überhaupt sinnvoll ist und, wenn ja, wie man sie praktisch am besten implementiert (für Übersichten vgl. Fox/Tracy 1986; Lensvelt-Mulders et al. 2005a, 2005b). Die Grundidee von RRT besteht darin, Anonymität bzw. Vertraulichkeit dadurch herzustellen, dass der Interviewer nicht weiß, worauf der Befragte antwortet, weil dies durch einen vom Befragten initiierten und kontrollierten Zufallsmechanismus bestimmt wird. Bei der RRT-Variante „forced response“ wirft die befragte Person z.B. zwei Münzen, und in Abhängigkeit vom Ergebnis des Münzwurfs, das der Interviewer nicht kennt, antwortet sie z.B. bei zweimal Wappen mit Ja (erzwungenes Ja), bei zweimal Zahl mit Nein (erzwungenes Nein), und nur bei einmal Wappen und einmal Zahl soll sie die heikle Frage beantworten. Andere RRT-Varianten arbeiten mit einer „unrelated question“ (z.B. ob die eigene Mutter an einem ungeraden oder einem geraden Tag geboren wurde), wobei wiederum mit einem Zufallsmechanismus bestimmt wird, ob die heikle oder die nicht-heikle Frage beantwortet wird. Da die von den

Interviewern registrierte Antwortverteilung lediglich mit einem bekannten Zufallsfehler behaftet ist, lässt sich mit einfachen Formeln die Prävalenz des heiklen Verhaltens schätzen (vgl. dazu die Übersichtstabelle bei Tourangeau/Yan 2007: 872).

Unsere Studierendenbefragung arbeitete mit „forced response“ mit Hilfe eines Kartenspiels. Alle Interviewer/innen, die indirekte Interviews durchführten, hatten ein Kartenspiel mit 12 Karten dabei (Karten im Format und Papiertyp von Skat-Karten). Auf 3 Karten stand ein großes Nein, auf 3 weiteren ein großes Ja und auf 6 Karten der Terminus „Wahrheit“. Die Interviewer überreichten den Befragten den Kartenstapel und erläuterten dessen Zusammensetzung. Nun sollten die Befragten vor jeder Frage, die der Interviewer stellte, den Kartenstapel eigenständig mischen und sich, ohne dem Interviewer Einblick zu gewähren, ansehen, welche Karte sie sich als unterste hingemischt hatten. Sodann stellte der Interviewer die Frage, und die Befragten antworteten mit Ja oder Nein, ohne dass der Interviewer wusste, ob es sich um ein erzwungenes oder ein wahres Ja/Nein handelte. Um zu testen, ob die Befragten die erläuterte Prozedur verstanden hatten, und zu Übungszwecken wurde mit einer „Justierungsfrage“ begonnen, die da lautete: „Wurde deine Mutter in den Monaten Oktober, November oder Dezember geboren: ja oder nein?“. Anschließend wurden neun Fragen zum „Uni-Schummeln“, d.h. zu unlauteren Praktiken beim Studium gestellt – immer zuerst mit Kartenmischen, dann die Frage, dann die Ja/Nein-Antwort. Tabelle 3 informiert darüber, welche Fragen in diesem Bereich gestellt wurden, und präsentiert die diesbezüglichen Ergebnisse mit dem interessierenden Vergleich von direktem und RRT-Fragemodus.⁸

Tabelle 3 hier

Eigentlich würde man erwarten, dass die RRT-Prozentwerte stets höher liegen als die entsprechenden Prozentwerte bei direktem Abfragen. Dies ist aber eindeutig nicht der Fall, denn alle drei möglichen Konstellationen ($RRT > \text{direkt}$, $RRT = \text{direkt}$, $RRT < \text{direkt}$) treten auf. Bei den zwei Items „Hausarbeit komplett plagiiert“ und „in Bibliothek Seiten aus einem Buch gerissen“ stehen in Tabelle 3 negative RRT-Prozentwerte. Dies ist zwar logisch unsinnig, kann aber speziell bei Verhaltensweisen mit geringer Prävalenz durchaus vorkommen, weil es sich bei den RRT-Prozentwerten ja um mit Zufallsfehlern behaftete Schätzungen handelt (ähnlich z.B. bei Coutts/Jann 2008). Folgt man den Signifikanztests in der letzten Spalte von Tabelle 3, zeigen sich bei fünf der neun Items keine signifikanten

⁸ Die RRT-Prozentwerte errechnen sich bei der von uns verwendeten „forced response/forced choice“-Methode nach der Formel: $p = (p_b - p_j) / p_s$, wobei p_b der Anteil derer ist, die in der Befragung die Antwort ja gegeben haben, p_j der Anteil derer, die durch den Zufallsmechanismus auf die Antwort ja gezwungen wurden, und p_s der Anteil derer, die auf die heikle Frage geantwortet haben. In unserem Kartenspiel waren $p_j = 0,25$ und $p_s = 0,50$.

Unterschiede, bei zwei Items ergeben sich signifikante Differenzen in Richtung der Ausgangshypothese, aber bei ebenfalls zwei Items erhält man signifikante Unterschiede entgegen der Ausgangshypothese.⁹

Auch die oben angesprochene und in Tabelle 3 ausgesparte RRT-„Justierungsfrage“, ob die eigene Mutter in den Monaten Oktober, November oder Dezember geboren wurde, hat nicht funktioniert. Denn es ergab sich ein RRT-Schätzwert von 40%, was von den erwarteten rund 22% (eigene Auszählung der Geburtsmonate von Frauen im Mikrozensus, die Monate November und Dezember sind „geburtenschwach“) deutlich abweicht.

Ebenfalls nicht in Tabelle 3 enthalten sind die zwei Ergänzungsfragen, ob die Beantwortung der Items zum „Uni-Schummeln“ unangenehm war und mehr oder weniger ehrlich erfolgt ist. Von den RRT-Befragten sagten 73%, dass diese Fragen für sie „überhaupt nicht unangenehm“ waren, von den direkt Befragten 79%, was von der Tendenz her eher gegen RRT spricht. Kaum Unterschiede zeigen sich bei der selbstbekundeten Ehrlichkeit in diesem Bereich, mit 88% „sehr ehrlich“ bei RRT und 92% bei der direkten Abfrage.

Was nun sind die Gründe für das Scheitern von RRT in der hier angewendeten Form? Vorab ist festzuhalten, dass RRT auch schon in anderen Studien ziemlich fragliche Ergebnisse gebracht hat (jüngst etwa bei Coutts/Jann 2008).¹⁰ In Nachbesprechungen zu unserer Erhebung wurde deutlich, dass die studentischen Interviewer/innen trotz vorheriger Schulung und Pretests RRT nicht ganz einheitlich umsetzten. Weiterhin wollten mehrere Befragte bei RRT offenbar einfach nicht mitmachen, weil sie die Frage lieber und aus eigener Sicht problemlos direkt beantworten wollten. Das Hauptproblem war aber eindeutig das erzwungene Ja. Wenn eine Person z.B. nie eine Hausarbeit plagiiert, sich aber eine Ja-Karte hingemischt hat, neigt sie dazu, mit dem Interviewer in eine Diskussion einzutreten und mitunter auf einem Nein zu beharren. Die „forced response“-Variante von RRT ist insofern etwas paradox, weil sie von ihrer Grundintention her unwahre Antworten verhindern will, genau zu diesem Zweck aber viele Befragte zu nicht-richtigen Antworten drängt (erzwungene Ja/Nein-Antworten). Die „unrelated question“-Variante von RRT hat dieses Problem nicht und erscheint von daher vorzugswürdig.

Ignoriert man die RRT-Werte in Tabelle 3, bleiben für die inhaltliche Interpretation immerhin

⁹ Wie in den Tabellen zuvor wird die Signifikanz des Unterschiedes über einseitige Chi-Quadrat-Tests ermittelt. Bei den zwei Items mit (unsinnigen) negativen RRT-Häufigkeiten wurden für die Signifikanztests die Werte null eingesetzt.

¹⁰ Dabei ist zu befürchten, dass zahlreiche „Fehlversuche“ mit RRT erst gar nicht zur Publikation gelangt sind.

noch die plausibleren direkten Werte. Demnach sind komplett gefälschte Hausarbeiten an der Universität sehr selten, nicht aber Teilfälschungen, die mit 16% zu Buche schlagen. Bei Uni-Klausuren sind das Abschreiben und das Abschreibenlassen offenbar weit verbreitete Praktiken. Demgegenüber spielen Unredlichkeiten bei der Bibliotheksnutzung so gut wie keine Rolle. Und bei Student/innen, die zumindest gelegentlich neben ihrem Studium jobben, ist Schwarzarbeit mit knapp 40% mehr als ein Randphänomen.

6. Schlussbemerkungen

Trotz gewisser Anfangsbedenken erwies sich der „sensitive topic survey“ bei unserer Zielgruppe der Studierenden als problemlos gangbar. Interviewabbrüche kamen so gut wie überhaupt nicht vor, die Item-Nonresponse-Raten lagen auch bei hochsensiblen Fragen (wie z.B. bei der Häufigkeit des Geschlechtsverkehrs in den letzten vier Wochen oder bei der Masturbationshäufigkeit) stets deutlich unter 10%,¹¹ und die Vorankündigung eines Interviews, in dem es inhaltlich um heikle Themen geht, hat bei etlichen Student/innen wohl erst das Interesse an einer Teilnahme geweckt.

Im Feld der vier Themenbereiche, auf die sich die heiklen Fragen bezogen (Alkohol/Drogen, kriminelle Alltagsdelikte, Sexualität, Uni-Schummeln), war aus der Sicht der Befragten der Sex-Teil eindeutig am heikelsten. Während die Sex-Fragen nur (oder vielleicht besser: immerhin) von rund einem Drittel der Interviewten als „überhaupt nicht unangenehm“ eingestuft wurden, waren es in den drei anderen Bereichen mehr als zwei Drittel. Folgt man den Angaben der Befragten, korreliert der Grad der Unangenehmheit der Fragen nur mäßig (negativ) mit der Ehrlichkeit der Antworten. Im Sex-Bereich haben angeblich rund 80% „sehr ehrlich“ geantwortet, in den drei anderen Bereichen rund 90%.

Die Techniken, die zur valideren Erfassung heikler Sachverhalte erprobt wurden, waren „Wording/Framing“, „vertrauliches Kuvert“ und „randomized response“. Die Ergebnisse bei Wording/Framing nähren den Verdacht, dass so manche Lehrbuchempfehlung, die von positiven Effekten euphemistischer Frageformulierungen ausgeht, einer stringenten empirischen Überprüfung nicht standhält. Demgegenüber ist der gelegentliche Einsatz vertraulicher Kuverts in face-to-face Interviews vermutlich vorteilhaft, insbesondere im Fall hochsensibler Fragen. RRT in der Variante von „forced response“ mit einem Kartenspiel, wie

¹¹ Man vergleiche dies mit den gängigen Verweigerungsquoten von 20-30% bei der Einkommensfrage. Tourangeau et al. (2000: 260ff.) belegen, dass die Höhe des Item-Nonresponse insgesamt kein guter Indikator für den Grad der „Heikelheit“ einer Frage ist.

wir es verwendet haben, erwies sich (drastisch formuliert) als ein Fehlschlag. Aber man würde wohl das Kind mit dem Bade ausschütten, dies sogleich zu verallgemeinern. Mit Sicherheit gilt, dass die Validität von RRT stark methodenabhängig ist. Schon kleine Änderungen im RRT-Design und bei der RRT-Implementation können erhebliche Auswirkungen haben und führen mitunter zu definitiv unbrauchbaren Ergebnissen.

Als eine Befragung von Studierenden an nur einer Universität (Mainz) hat die vorliegende Studie von der Stichprobe her eine klare Begrenzung. Die fragliche Verallgemeinerbarkeit auf die Gesamtbevölkerung ist aber nicht unbedingt problematisch, weil es ja vor allem um einen Methodenvergleich ging (Vergleich von direktem und indirektem Fragemodus). Auch das Spektrum der Techniken, die zur Erhebung heikler Sachverhalte zum Einsatz gelangten, war begrenzt und ist zweifellos erweiterungsfähig (vgl. dazu die Hinweise im einleitenden Abschnitt 1). Schließlich sei vermerkt, dass in unserer Untersuchung vom begrenzten „klassischen Konzept“ heikler Fragen ausgegangen wurde, d.h. von Verhaltensfragen mit einer Tendenz zu Underreporting (Tourangeau et al. 2000: 269ff.). Das Forschungsfeld erweitert sich, wenn man auch Verhaltensfragen mit einer Tendenz zu Overreporting einbezieht (z.B. Wahlbeteiligung oder Kirchgangshäufigkeit), und es erweitert sich noch einmal, wenn man auch heikle Einstellungen einschließt (z.B. autoritäre Einstellungen oder Diskriminierungsneigungen). Ob letzteres sinnvoll ist, ist vor allem deshalb umstritten, weil dann der für Validierungsstudien wichtige Referenzpunkt eines „wahren Wertes“ verloren geht.

Literatur

Barnett, Julie, 1998: Sensitive Questions and Response Effects: An Evaluation. *Journal of Managerial Psychology* 13: 63-76.

Barton, Allen H., 1958: Asking the Embarrassing Question. *Public Opinion Quarterly* 22: 67-68.

Becker, Rolf und Ralph Günther, 2004: Selektives Antwortverhalten bei Fragen zum delinquenten Handeln. Eine empirische Studie über die Wirksamkeit der „sealed envelope technique“ bei selbst berichteter Delinquenz mit Daten des ALLBUS 2000. *ZUMA-Nachrichten* 28 (Nr. 54): 39-59.

Bradburn, Norman M. und Seymour Sudman, 1979: *Improving Interview Method and Questionnaire Design*. San Francisco: Jossey-Bass.

Coutts, Elisabeth und Ben Jann, 2008: Sensitive Questions in Online Surveys: Experimental Results for the Randomized Response Technique (RRT) and the Unmatched Count Technique (UCT). Zürich: ETH-Zürich Sociology Working Paper No. 3.

Deffaa, Walter, 1982: *Anonymisierte Befragungen mit zufallsverschlüsselten Antworten*. Frankfurt am Main: Peter Lang.

De Leeuw, Edith D., 2001: Reducing Missing Data in Surveys: An Overview of Methods. *Quality and Quantity* 35: 147-160.

Diekmann, Andreas, 2007: *Empirische Sozialforschung*. 18. Auflage. Reinbek bei Hamburg: Rowohlt.

Fox, James Alan und Paul E. Tracy, 1986: *Randomized Response: A Method for Sensitive Surveys*. Newbury Park: Sage.

Hox, Joop J. und Edith D. de Leeuw, 2002: The Influence of Interviewers' Attitude and Behavior on Household Survey Nonresponse: An International Comparison. S. 103-120 in: Robert M. Groves, Don A. Dillman, John L. Eltinge und Roderick J.A. Little (Eds.): *Survey Nonresponse*. New York: Wiley.

Jasso, Guillermina, 1985: Marital Coital Frequency and the Passage of Time: Estimating the Separate Effects of Spouses' Ages and Marital Duration, Birth and Marriage Cohorts, and Period Influences. *American Sociological Review* 50: 224-241.

Lee, Raymond M., 1993: *Doing Research on Sensitive Topics*. London: Sage.

Lensvelt-Mulders, Gerty J.L.M., Joop J. Hox, Peter G.M. van der Heijden und Cora J.M. Mass, 2005a: Meta-Analysis of Randomized Response Research. Thirty-Five Years of Validation. *Sociological Methods & Research* 33: 319-348.

Lensvelt-Mulders, Gerty J.L.M., Joop J. Hox und Peter G.M. van der Heijden, 2005b: How to

Improve the Efficiency of Randomized Response Designs. *Quality and Quantity* 39: 253-265.

Renzetti, Claire M. und Raymond M. Lee (Eds.), 1993: *Researching Sensitive Topics*. Newbury Park: Sage.

Schnell, Rainer, 1997: *Nonresponse in Bevölkerungsumfragen*. Opladen: Leske und Budrich.

Schnell, Rainer, Paul B. Hill und Elke Esser, 2005: *Methoden der empirischen Sozialforschung*. 7. Auflage. München: Oldenbourg.

Schnell, Rainer und Frauke Kreuter, 2005: Separating Interviewer and Sampling-Point Effects. *Journal of Official Statistics* 21: 389-410.

Sudman, Seymour und Norman M. Bradburn, 1974: *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine.

Sudman, Seymour und Norman M. Bradburn, 1982: *Asking Questions: A Practical Guide to Questionnaire Design*. San Francisco: Jossey-Bass.

Tourangeau, Roger, Lance J. Rips und Kenneth A. Rasinski, 2000: *The Psychology of Survey Response*. Cambridge: Cambridge University Press.

Tourangeau, Roger und Ting Yan, 2007: Sensitive Questions in Surveys. *Psychological Bulletin* 133: 839-883.

Van Koolwijk, Jürgen, 1969: Unangenehme Fragen. Paradigma für die Reaktionen des Befragten im Interview. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 21: 864-875.

Warner, Stanley L., 1965: Randomized Response: A Survey Technique for Eliminating Evasive Answer Bias. *Journal of the American Statistical Association* 60: 63-69.

Tabelle 1: Wording/Framing-Techniken (WFT) im Vergleich zu direkten Fragen (Prozentangaben)

	Direkt (n=211)	WFT (n=367)	Signifikanz des Unterschiedes
Alkoholkonsum: täglich, 4-6x pro Woche	6,6	7,9	p=0,58
Vollrausch: sehr häufig, häufig, gelegentlich	28,0	28,9	p=0,81
Illegale Drogen: jemals	54,8	60,5	p=0,18
GEZ-Betrug: aktuell	29,8	39,2	p=0,04*
Bußgeld: jemals	66,4	66,4	p=0,99
Ladendiebstahl: jemals	29,4	32,2	p=0,48
Strafrechtliches Verfahren: jemals	8,5	11,3	p=0,30

Alkoholkonsum: 1) direkte Frage „Wie häufig trinkst du Alkohol?“ mit den sechs Antwortvorgaben täglich, 4-6 mal pro Woche, 2-3 mal pro Woche, 1 mal pro Woche, seltener, nie, 2) WFT-Frage „Zahlreiche empirische Studien kommen zu dem Ergebnis, dass mit mäßigem Alkoholkonsum (z.B. mit einem Glas Rotwein oder mit einem abendlichen Bier) auch positive Effekte einher gehen. Wie häufig trinkst du Alkohol?“.

Vollrausch: 1) direkte Frage „Und wie häufig kommt es vor, dass du so richtig betrunken bist?“ mit den fünf Antwortvorgaben sehr häufig, häufig, gelegentlich, selten, nie, 2) WFT-Frage „Betrunken ist ja jeder mal! Und manchmal gehört es einfach dazu, auch einen über den Durst zu trinken. Wie häufig kommt es vor, dass du so richtig betrunken bist?“.

Illegale Drogen: 1) direkte Frage „Hast du schon jemals illegale Drogen genommen?“ mit den Antwortvorgaben ja/nein, 2) WFT-Frage „Auch VIPs, also sehr bedeutsame Personen aus dem öffentlichen Leben berichten, in ihrer Jugend sogenannte illegale Drogen (angefangen von Marihuana bis hin zu härteren Sachen) probiert und genommen zu haben. Hast du schon jemals illegale Drogen genommen?“.

GEZ-Betrug: 1) direkte Frage „Sicher kennst du die GEZ, die Gebühren-Einzugs-Zentrale in Köln, die für die Radio- und Fernsehgebühren zuständig ist. Zahlst du oder dein Haushalt, in dem du lebst, zur Zeit Gebühren an die GEZ: ja oder nein?“, 2) der Wortlaut der WFT-Frage war wie bei der direkten Frage, aber nach der Trichter-Technik waren mehrere Fragen hinführend vorgeschaltet, Fragen nach der Wohnform, nach dem Vorhandensein eines Fernsehers/Radios/Computers im Haushalt und nach der Beurteilung des Tatbestandes, dass viele Studierende keine GEZ-Gebühren zahlen, obwohl sie es eigentlich tun müssten (finde ich voll okay bis überhaupt nicht okay).

Bußgeld: 1) direkte Frage „Hast du jemals ein Buß- oder Verwarnungsgeld für eine Verkehrsordnungswidrigkeit zahlen müssen?“ mit den Antwortvorgaben ja/nein, 2) der Wortlaut der WFT-Frage war wie bei der direkten Frage, aber

nach der Trichter-Technik waren mehrere Fragen hinführend vorgeschaltet, Fragen nach der Hektik im Stadtverkehr, nach der Intensität des eigenen Bemühens um Einhaltung der Straßenverkehrsordnung, nach der Nutzungshäufigkeit verschiedener Verkehrsmittel und nach der Einschätzung der Wichtigkeit der Befolgung verschiedener Verkehrsregeln (rote Ampeln, Tempolimits usw.).

Ladendiebstahl: 1) direkte Frage „Hast du schon einmal einen Ladendiebstahl begangen?“ mit den Antwortvorgaben ja, nein, weiß nicht mehr, 2) der Wortlaut der WFT-Frage war wie bei der direkten Frage, aber nach der Trichter-Technik waren mehrere Fragen hinführend vorgeschaltet, Fragen nach unangemessenen Preiserhöhungen im Zuge der EURO-Umstellung, nach der Vertretbarkeit eines Ladendiebstahls unter bestimmten Bedingungen (materielle Notlage) und nach der mutmaßlichen Häufigkeit von Ladendiebstählen.

Strafrechtliches Verfahren: 1) direkte Frage „Wurde aus irgendeinem Grund gegen dich jemals strafrechtlich ermittelt? Strafrechtlich bedeutet, dass die Sache über den Staatsanwalt gelaufen ist?“ mit den Antwortvorgaben ja, nein, weiß nicht, 2) der Wortlaut der WFT-Frage war wie bei der direkten Frage, aber nach der Trichter-Technik waren mehrere Fragen hinführend vorgeschaltet, Fragen nach der Angst vor Kriminalität, nach der Beobachtung von kriminellen Delikten (Augenzeuge einer Straftat), nach eigener Betroffenheit von Kriminalität (Viktimisierung) und nach dem Dabeisein bei Gerichtsverhandlungen.

Tabelle 2: Technik des vertraulichen Kuverts (VKT) im Vergleich zu direkten Fragen (Mittelwerte und Prozentangaben)

		Direkt (n=211)	VKT (n=367)	Signifikanz des Unterschiedes
Sexualpartner: bisherige Zahl	Median: Mittel:	4,0 5,9	4,0 6,6	p=0,40
Geschlechtsverkehr: Häufigkeit in den letzten 4 Wochen	Median: Mittel:	3,0 4,7	3,0 5,0	p=0,62
One-Night-Stand: jemals		53,8	49,3	p=0,30
Seitensprung: in aktueller oder letzter Beziehung		12,5	10,5	p=0,48
Homosexuelle Kontakte: jemals		14,8	18,5	p=0,26
Selbstbefriedigung: häufig, sehr häufig		14,6	29,4	p=0,00*

Sexualpartner: Frage „Mit wie vielen unterschiedlichen Partner/innen hattest du in deinem bisherigen Leben Sex? Es genügt, wenn du gegebenenfalls eine Schätzung der Zahl der Partner/innen gibst“.

Geschlechtsverkehr: Frage „In den letzten vier Wochen: Wieviel Mal hattest du da Geschlechtsverkehr? Es genügt, wenn du gegebenenfalls eine Schätzung gibst“.

One-Night-Stand: Frage „Hattest du in deinem Leben schon einmal einen so genannten One-Night-Stand?“ mit den Antwortvorgaben ja/nein.

Seitensprung: Frage „In deiner aktuellen (bzw. letzten) festen Beziehung: Hattest du da schon einmal einen Seitensprung? Mit Seitensprung ist Sex mit einer anderen Person gemeint, wenn jemand gleichzeitig eine feste Beziehung hat“ mit den Antwortvorgaben ja/nein.

Homosexuelle Kontakte: Frage „Hattest du selbst in deinem bisherigen Leben jemals homosexuelle Kontakte?“ mit den Antwortvorgaben ja/nein.

Selbstbefriedigung: Frage „Obwohl nach wissenschaftlichen Studien mehr als 80% der Männer und mehr als 60% der Frauen sich selbst befriedigen, ist in der Sexualkunde und mehr noch im Alltagsleben Selbstbefriedigung das am stärksten tabuisierte Thema. Trotzdem erlauben wir uns die Frage: Wie häufig befriedigst du dich selbst?“ mit den Antwortvorgaben sehr häufig, häufig, gelegentlich, selten, nie.

Tabelle 3: Randomized-Response-Technik (RRT) im Vergleich zu direkten Fragen (Prozentangaben)

	Direkt (n=211)	RRT (n=367)	Signifikanz des Unterschiedes
Hausarbeit: komplett plagiiert	1,4	-2,5	p=0,29
Hausarbeit: teilweise plagiiert	15,9	4,9	p=0,00*
Klausur: abschreiben lassen	65,7	63,1	p=0,60
Klausur: selbst abgeschrieben	57,0	41,1	p=0,00*
Klausur: Spickzettel u.ä. verwendet	21,7	43,9	p=0,00*
Bibliothek: Buch versteckt	5,4	4,4	p=0,71
Bibliothek: Buch entwendet	0,0	6,7	p=0,00*
Bibliothek: Seiten aus Buch gerissen	1,0	-4,7	p=0,29
Nebenjob: schwarz gearbeitet	39,4	41,4	p=0,41

Hausarbeit komplett plagiiert: Frage „Hast du in deiner Uni-Zeit schon einmal eine Hausarbeit komplett aus dem Internet übernommen, ohne die Quelle anzugeben?“ mit den Antwortvorgaben ja/nein.

Hausarbeit teilweise plagiiert: Frage „Hast du in deiner Uni-Zeit schon einmal eine Hausarbeit teilweise aus dem Internet übernommen, ohne die Quelle anzugeben?“ mit den Antwortvorgaben ja/nein.

Klausur abschreiben lassen: Frage „Hast du in deiner Uni-Zeit schon einmal bei einer Klausur bewusst zugelassen, dass eine Kommilitonin oder ein Kommilitone bei dir abschreibt?“ mit den Antwortvorgaben ja/nein.

Klausur selbst abgeschrieben: Frage „Und umgekehrt: Hast du in deiner Uni-Zeit schon einmal bei einer Klausur selbst von einer Kommilitonin oder einem Kommilitonen abgeschrieben?“ mit den Antwortvorgaben ja/nein.

Klausur Spickzettel u.ä. verwendet: Frage „Hast du in deiner Uni-Zeit schon einmal bei einer Klausur einen Spickzettel oder andere unerlaubte Materialien verwendet?“ mit den Antwortvorgaben ja/nein.

Bibliothek Buch versteckt: Frage „Hast du an der Uni schon jemals in der Bibliothek ein oder mehrere Bücher versteckt?“ mit den Antwortvorgaben ja/nein.

Bibliothek Buch entwendet: Frage „Hast du an der Uni schon jemals in der Bibliothek ein Buch entwendet?“ mit den Antwortvorgaben ja/nein.

Bibliothek Seiten aus Buch gerissen: Frage „Und hast du an der Uni schon jemals in der Bibliothek Seiten aus einem Buch herausgerissen?“ mit den Antwortvorgaben ja/nein.

Nebenjob schwarz gearbeitet: Frage „Bei dem oder den Jobs neben deinem Studium: Hast du da schon einmal schwarzgearbeitet?“ mit den Antwortvorgaben ja/nein.