

Themenschwerpunkt

Studiendesign und Methodik des Epidemiologischen Suchtsurveys 2009

Study Design and Methodology of the 2009 Epidemiological Survey of Substance Abuse

Ludwig Kraus und Alexander Pabst

IFT Institut für Therapieforchung, München

Zusammenfassung: *Hintergrund:* Der Epidemiologische Suchtsurvey (ESA) 2009 untersucht den Konsum und Missbrauch von psychoaktiven Substanzen in der deutschen Allgemeinbevölkerung. *Ziel:* Beschreibung der Stichprobenziehung, der eingesetzten Instrumente, der Datenerhebung, Ausschöpfung und Gewichtung des Surveys. Durchführung einer Nonresponse- und Modus-Effekt-Analyse. *Methodik:* Die Stichprobe wurde disproportional zur Verteilung der Geburtskohorten in der Bevölkerung in einem zweistufigen, zufalls-gesteuerten Auswahlverfahren aus den Einwohnermelderegistern gezogen. *Ergebnisse:* Insgesamt nahmen 8.030 Personen (3.731 schriftlich, 3.376 telefonisch und 927 online) im Alter zwischen 18 und 64 Jahren teil. Die Antwortrate betrug 50.1 %. Die verwendete Design- und Poststratifikationsgewichtung erzielte eine Effektivität von 83.5 %. Abweichungen von der Grundgesamtheit bestehen in der gewichteten Stichprobe hinsichtlich Staatsbürgerschaft und Schulbildung. An der telefonischen oder schriftlichen Kurzbefragung nahmen 8 % aller Nichtantworter teil. Die Vergleiche zeigen, dass Nichtteilnehmer seltener Alkohol- und Drogenkonsum berichteten als Teilnehmer. Modus-Effekte deuten im Vergleich zu schriftlich Befragten auf eine geringere Verbreitung von Cannabiskonsum unter telefonisch Befragten und eine geringere Prävalenz von Alkoholkonsum unter Online-Teilnehmern hin. *Schlussfolgerungen:* Der eingesetzte Methodenmix aus schriftlicher, telefonischer und Online-Befragung wirkte sich positiv auf die Teilnahmebereitschaft aus, was sich in einer gestiegenen Antwortrate im Vergleich zu früheren ESA-Studien zeigt.

Schlüsselwörter: Epidemiologie, Bevölkerungsbefragung, Methodik, Instrumente, Nonresponder, Validität

Abstract: *Background:* The 2009 Epidemiological Survey of Substance Abuse (ESA) investigates intake and abuse of psychoactive substances in the German general population. *Aims:* To describe sampling, instruments, data collection, response rate, weighting procedure, and to perform non-response and mode-effect analyses. *Method:* Using a two-stage probability sampling design, subjects were randomly selected from population registers, with younger birth cohorts being over- and older birth cohorts undersampled. *Results:* The full sample comprised 8,030 respondents aged 18 to 64 years. Subjects participated either by postal questionnaire ($n = 3,731$), telephone interview ($n = 3,376$) or online ($n = 927$). The estimated response rate was 50.1 %. The design- and poststratification weighting achieved an effectiveness of 83.5 %. Deviations from the population distributions in the weighted sample exist with regard to citizenship and education. Brief telephone interviews or postal questionnaires were administered to all nonparticipants with 8 % responding. Nonresponse analyses revealed that nonparticipants reported less alcohol and drug use than participants. Subjects who responded by telephone reported less cannabis use, and subjects who participated online reported less alcohol use than those who answered the postal questionnaire. *Conclusions:* The data collection procedure including postal, telephone, and online administration positively affected the willingness to participate, resulting in a higher response rate as compared to earlier ESA studies.

Keywords: epidemiology, population survey, methodology, instruments, nonresponse, validity

Einleitung

Surveys sind für die epidemiologische Forschung das Mittel der Wahl zur Gewinnung von Daten zu Konsum und Konsummustern psychoaktiver Substanzen in Subgruppen oder der ganzen Bevölkerung. Der regelmäßig bundesweit durchgeführte Epidemiologische Suchtsurvey (ESA) trägt

damit wesentlich zur Beschreibung und zum Verständnis von Substanzkonsum und süchtigem Verhalten bei und leistet einen wichtigen Beitrag zur Verbesserung der öffentlichen Gesundheit und zur Prävention.

Die ursprünglich im Jugend- und Erwachsenenbereich angesiedelte Studie wurde erstmals 1980 auf dem Gebiet der damaligen Bundesrepublik Deutschland durchgeführt.

Nach der Wiedervereinigung wurde der Survey auf ganz Deutschland ausgedehnt und auf eine reine Erwachsenenbefragung im Altersbereich von 18 bis 59 Jahren umgestellt. In der Erhebung von 2006 wurde der Altersbereich bis 64 Jahre erweitert. Mit dem Survey von 2009 liegt die zehnte Erhebung in der erwachsenen deutschen Allgemeinbevölkerung vor.

Neben aktuellen epidemiologischen Daten zum Konsum und Missbrauch psychoaktiver Substanzen liefert der Epidemiologische Suchtsurvey wertvolle Informationen über abhängiges Verhalten und problematische Konsummuster unterhalb der Schwelle von Missbrauch und Abhängigkeit. Als wiederholte Querschnittsbefragung erlaubt der ESA zudem die Beschreibung von zeitlichen Veränderungen im Konsumverhalten (Baumeister, Kraus, Stonner & Metz, 2008). Weiterhin eignen sich die Daten für die Untersuchung spezifischer Fragestellungen wie dem Einfluss von Präventionsmaßnahmen auf das Konsumverhalten (Müller, Piontek, Pabst, Baumeister & Kraus, 2010), den Zusammenhang zwischen Alkoholwirkungserwartungen und Konsum (Pabst, Baumeister & Kraus, 2010) oder die Analyse negativer Auswirkungen exzessiven Konsumverhaltens (Kraus, Baumeister, Pabst & Orth, 2009).

Das Ziel dieser Arbeit ist es, die Methodik des Epidemiologischen Suchtsurveys 2009 (Stichprobenziehung, Datenerhebung und Ausschöpfung) zu beschreiben. Im Rahmen des Qualitätsmonitorings sollen dafür die Güte der verwendeten Gewichtung sowie systematische Verzerrungen der realisierten Stichprobe untersucht werden. Hierfür werden die Teilnehmer an der Studie mit Nichtteilnehmern verglichen, die bereit waren, sechs kurze Fragen zu ihrem Konsumverhalten zu beantworten. Weiterhin werden potentielle Modus-Effekte zwischen telefonisch, schriftlich und im Internet Befragten untersucht, um daraus Aussagen über mögliche Selektionseffekte der erzielten Stichprobe ableiten zu können.

Methoden

Stichprobe

Die Grundgesamtheit des Epidemiologischen Suchtsurveys 2009 umfasst alle in Privathaushalten lebenden deutschsprachigen Personen, die in den Jahren zwischen 1945 und 1991 geboren wurden. Die Stichprobe basiert somit auf etwa 51.6 Mio. Personen der deutschen Allgemeinbevölkerung im Alter zwischen 18 und 64 Jahren (Stand 31.12.2008, Statistisches Bundesamt). Die Stichprobenziehung erfolgte in einem zweistufigen zufallsgesteuerten Auswahlverfahren. Auf der ersten Stufe wurden Schichtungszellen aus der Kombination von Landkreisen, Regierungsbezirken und Bundesländern mit zehn BIK-Gemeindegrößenklassen gebildet. Die Gemeindegrößenklassen kategorisieren die Gemeinden und Stadtteile nach Einwohnerzahl und Kernstadt bzw. Randregion. Innerhalb dieser

Schichtungszellen wurden 238 Primäreinheiten bestehend aus Gemeinden bzw. Stadtteilen in Großstädten proportional zur Anzahl der in ihnen lebenden Personen zufällig ausgewählt. Dies gewährleistete die größtmögliche geographische und siedlungsstrukturelle Proportionalität der Stichprobe zur Grundgesamtheit.

Die Ziehung der Personenadressen in den ausgewählten Primäreinheiten erfolgte auf der zweiten Stufe über eine systematische Zufallsauswahl aus den Melderegistern der Einwohnermeldeämter. Je Primäreinheit wurde dabei die gleiche Zahl von zu befragenden Personen eingesetzt. Um den geringeren Anteil junger Erwachsener an der Gesamtbevölkerung auszugleichen, wurden die Personenadressen in den ausgewählten Primäreinheiten nach Geschlecht und Geburtsjahrgang (1945–49, 1950–59, 1960–69, 1970–79, 1980–84, 1985–88, 1989–91) stratifiziert und die benötigten Zieladressen je Primäreinheit disproportional zur Verteilung in der Grundgesamtheit gezogen. Damit wurde gewährleistet, dass in der Stichprobe alle Altersgruppen in ausreichender Fallzahl vertreten waren. Insgesamt wurden 8.000 Fälle für die Nettostichprobe angestrebt mit jeweils 1.000 Personen in den drei jüngsten (1980–91), 1.400 in den drei folgenden (1950–79) und 800 Personen in der ältesten Jahrgangsguppe (1945–49).

Instrumente

Der Fragebogen des ESA 2009 stellt eine adaptierte Version des 2006 eingesetzten Instrumentes dar (Kraus & Baumeister, 2008). Neben dem regelmäßigen Monitoring des Substanzkonsums wurden neue Untersuchungsschwerpunkte gesetzt und auf aktuelle Forschungsfragen eingegangen. Im Zentrum stand wie in den vorangegangenen Jahren die Untersuchung von Konsumverhalten und Störungen im Zusammenhang mit dem Gebrauch von illegalen Substanzen, Alkohol, Tabak und Medikamenten. Neben soziodemographischen Daten und Fragen zur ökonomischen Situation wurden zudem der körperliche und psychische Gesundheitszustand sowie pathologisches Spielverhalten erfasst. Die Erhebung soziodemographischer Merkmale wurde dabei basierend auf den Demographischen Standards von 2004 (Statistisches Bundesamt, 2004) überarbeitet und um Fragen zum Migrationshintergrund erweitert.

In der vorliegenden Erhebung wurde der Fragebogen um zusätzliche Fragemodule zu Kognitionen im Zusammenhang mit Alkoholkonsum ergänzt. Erfasst wurden dabei neben trinkbezogener Selbstwirksamkeit (Oei, Hasking & Young, 2005) auch Alkoholwirkungserwartungen (Hagen & Demmel, 2002). Bei Fragen zum Alkoholkonsum wurde auf die Erfassung von Alkopops als separate Getränkeorte verzichtet, da die Verbreitung dieser Getränke in den letzten Jahren deutlich zurückgegangen ist (Bundeszentrale für gesundheitliche Aufklärung, 2007, 2008). Neu erfragt wurden stattdessen alkoholhaltige Mixgetränke als spezifische Getränkeorte. Die Fragen zum Konsum illegaler Substan-

zen wurden um synthetische Cannabinoide (Spice, Smoke, Space o. a.) erweitert. Zur Erfassung substanzbezogener Störungen und problematischen Konsums wurden Screeninginstrumente verwendet: die Severity of Dependence Skala (SDS; Gossop et al., 1995) für Cannabis-, Kokain- und Amphetaminabhängigkeit, der Alcohol Use Disorders Identification Test (AUDIT; Babor, Higgins-Biddle, Saunders & Monteiro, 2001) für problematischen Alkoholkonsum sowie der Kurzfragebogen zum Medikamentengebrauch (KFM; Watzl, Rist, Höcker & Miehle, 1991) für problematischen Medikamentengebrauch. Die SDS (Cannabis) wurde bereits im ESA 2006 verwendet, der AUDIT im ESA 1997 und 2003, und der KFM in allen Erhebungen seit dem Jahr 2000.

Aus Platzgründen wurde der Fragebogen an anderer Stelle gekürzt. Verzichtet wurde auf Fragen zur wahrgenommenen sozialen Unterstützung (Franke, Mohn, Sitzler, Wellbrink & Witte, 2001), zur Änderungsbereitschaft von Tabakkonsumenten (Etter & Perneger, 1999), zu Gewohnheiten und Verhaltensweisen im Zusammenhang mit Tabakkonsum (Fagerström-Test; Heatherton, Kozlowski, Frecker & Fagerström, 1991) und zu medikamentenorientierten Einstellungen (Schär, Hornung, Gutscher & May, 2007). Da Änderungen der Prävalenz substanzbezogener Störungen über kurze Zeiträume nicht zu erwarten sind, wurde im Survey 2009 mit Ausnahme von Nikotinabhängigkeit auf die Erfassung der umfangreichen DSM-IV Kriterien (American Psychiatric Association, 1994) zugunsten der kürzeren Screeninginstrumente SDS, AUDIT und KFM verzichtet. Die Erfassung von Nikotinabhängigkeit wurde beibehalten, um den Einfluss des zwischen den Erhebungen in Kraft getretenen Nichtraucherschutzgesetzes zu untersuchen.

Häufigkeit und Menge des Substanzkonsums

Konsum illegaler Drogen

Die Prävalenz und die Häufigkeit des Konsums illegaler Drogen wurde bezogen auf die Lebenszeit, die letzten 12 Monate und die letzten 30 Tage vor der Befragung für die Substanzen Cannabis, Amphetamine, Ecstasy, LSD, Heroin, andere Opiate, Kokain, Crack, Pilze und Spice erfragt. Zusätzlich wurde das Erstkonsumalter dieser Substanzen erhoben.

Alkoholkonsum

Alkoholkonsum in den letzten 12 Monaten und den letzten 30 Tagen vor der Befragung wurde mittels eines getränkespezifischen Frequenz-Menge-Index erhoben (Gmel & Rehm, 2004). Neben Bier, Wein/Sekt und Spirituosen wurden dabei erstmals alkoholhaltige Mixgetränke (Alkopops, Longdrinks (z. B. Wodka-Lemon), Cocktails (z. B. Caipirinha) oder Bowle) als Getränkearten erfasst. Für jedes Getränk wurden die Häufigkeit des Konsums sowie die durch-

schnittliche Trinkmenge an einem Konsumtag erfragt und mittels getränkespezifischer Alkoholgehalte (in Vol %; Bühringer et al., 2000) der durchschnittliche Tageskonsum berechnet. Zusätzlich wurde die Häufigkeit des Konsums von fünf oder mehr Gläsern Alkohol zu einer Trinkgelegenheit als Indikator für episodisches Rauschtrinken erhoben (Gmel, Rehm & Kuntsche, 2003).

Tabakkonsum

Als Raucher wurden Personen definiert, die in den letzten 30 Tagen vor der Befragung Zigaretten, Zigarren, Zigarillos oder Pfeife geraucht haben. Nichtraucher waren Personen, die insgesamt im Leben nicht mehr als 100 Zigaretten (bzw. andere Tabakprodukte) geraucht haben. Als Ex-Raucher galten Personen, die schon mehr als 100 Zigaretten oder andere Tabakprodukte konsumiert haben, in den letzten 30 Tagen jedoch abstinent waren. Die durchschnittliche Zigarettenzahl pro Tag wurde analog zu Alkoholkonsum mit einem Frequenz-Menge-Index für die letzten 30 Tage erhoben und ergibt sich aus der Anzahl der Rauchtage und der durchschnittlichen Anzahl der an einem Tag konsumierten Zigaretten. Starke Raucher wurden als Raucher von täglich 20 Zigaretten oder mehr definiert.

Medikamentengebrauch

Für den Zeitraum der letzten 12 Monate und der letzten 30 Tage vor der Befragung wurde die Prävalenz des Gebrauchs von Schmerzmitteln (Analgetika), Schlafmitteln (Hypnotika), Beruhigungsmitteln (Tranquilizer), Anregungsmitteln (Analeptika) und Appetitzüglern (Anorektika) erfragt. Vor dem Hintergrund häufiger Fehlanwendungen wurden zusätzlich Antidepressiva und Neuroleptika abgefragt. Um die Zuordnung der konsumierten Medikamente zu den Arzneimittelgruppen durch die Befragten zu erleichtern, wurde eine Liste mit den gebräuchlichsten Präparaten und deren Zuordnung zu den genannten Gruppen vorangestellt. Zusätzlich zur Prävalenz wurde die Häufigkeit des Medikamentengebrauchs in den letzten 30 Tagen erhoben.

Glücksspielverhalten

Das Spielverhalten wurde für verschiedene Glücksspiele bezogen auf die Lebenszeit, die letzten 12 Monate und die Häufigkeit in den letzten 30 Tagen erfasst. Die erfragten Glücksspiele umfassen die Gruppen Fernsehlotterie, Klassenlotterie, Lotto/Toto/Keno/Zusatzlotterie, Quicky, Lose, Sportwetten in Annahmestellen und im Internet, Pferdewetten, Geldspielautomaten, großes und kleines Spiel im Spielkasino, Pokerturniere bzw. Kartenspiele im Internet und illegales Glücksspiel.

Substanzbezogene Störungen

Drogenbezogene Störungen

Ein Screening für das Vorliegen einer Abhängigkeit von Cannabis, Amphetaminen oder Kokain wurde mit der Severity of Dependence Skala (SDS; Gossop et al., 1995; Steiner, Baumeister & Kraus, 2008) für den Zeitraum der letzten 12 Monate durchgeführt. Die SDS erfasst das psychologische Abhängigkeitserleben von Konsumenten illegaler Substanzen und enthält fünf Items zu Kontrollverlust, Kompulsivität des Gebrauchs, Besorgnis über den Konsum sowie zum Aufhörwunsch.

Problematischer Alkoholkonsum

Problematischer Alkoholkonsum in den letzten 12 Monaten wurde mit dem Alcohol Use Disorders Identification Test (AUDIT; Babor et al., 2001; deutschsprachige Version AUDIT-GM; Rist, Scheuren, Demmel, Hagen & Aulhorn, 2003) erfasst. Der AUDIT kann zum Screening von Alkoholabhängigkeit sowie zur Früherkennung alkoholbedingter Störungen bei Personen mit riskantem Alkoholkonsum vor dem Auftreten von physischen, psychischen oder sozialen Folgeschäden eingesetzt werden. Die 10 Items umfassen Fragen zu Menge und Frequenz des Alkoholkonsums sowie zu Symptomen und Konsequenzen von Alkoholmissbrauch und Alkoholabhängigkeit.

Nikotinabhängigkeit

Nikotinabhängigkeit wurde für den Zeitraum der letzten 12 Monate nach den Kriterien des DSM-IV (American Psychiatric Association, 1994) mit den Items des Münchener Composite International Diagnostic Interview (M-CIDI; Lachner et al., 1998; Wittchen et al., 1995) erfasst.

Problematischer Medikamentengebrauch

Zum Screening problematischen Medikamentengebrauchs in den letzten 12 Monaten wurde der Kurzfragebogen zum Medikamentengebrauch (KFM; Watzl et al., 1991) eingesetzt. Der aus 11 Items bestehende KFM erfasst verhaltensbezogene und emotionale Aspekte der Medikamenteneinnahme und dient der Vororientierung in Bezug auf einen möglichen Missbrauch bzw. eine Abhängigkeit von Medikamenten.

Pathologisches Glücksspiel

Pathologisches Glücksspiel in den letzten 12 Monaten wurde nach den Kriterien des DSM-IV (American Psychiatric Association, 1994) mit den ins Deutsche übersetzten Items von Stinchfield (2002) erfasst. Die Abfrage erfolgte für Personen, die angaben, im vergangenen Jahr im Durchschnitt mehr als 50 EUR pro Monat für Glücksspiele eingesetzt zu haben. Einsätze unterhalb dieses Grenzwertes

werden als geringes Risiko für pathologisches Glücksspiel erachtet (Currie et al., 2006). Zum Ausschluss einer manischen Episode gemäß DSM-IV wurde eine Screeningfrage zur Maniesymptomatik verwendet (siehe Abschnitt Psychische Gesundheit).

Kognitive Einflussfaktoren auf Alkoholkonsum

Alkoholwirkungserwartungen

Alkoholwirkungserwartungen kennzeichnen subjektive Erwartungen an die unmittelbare Wirkung von Alkohol auf das eigene Erleben und Verhalten und wurden mit den 20 Items des Inventars akuter Alkoholeffekte (IAA; Hagen & Demmel, 2002; Hagen, Demmel & Rist, 2002) erhoben. Erfasst wurden sowohl positive (z. B. Spannungsreduktion) als auch negative (z. B. Aggressionssteigerung) Wirkungserwartungen, die mit Alkoholkonsum assoziiert sind. Die Probanden sollten dabei die Wahrscheinlichkeit einschätzen, mit der sie erwarten, dass Alkohol bei ihnen in der jeweils beschriebenen Art und Weise wirkt.

Trinkbezogene Selbstwirksamkeitserwartungen

Selbstwirksamkeit (Bandura, 1997) im Kontext von Alkoholkonsum bezieht sich auf die Einschätzung der eigenen Fähigkeit, dem Konsum alkoholischer Getränke zu widerstehen. Solche trinkbezogenen Selbstwirksamkeitserwartungen wurden mit der ins Deutsche übersetzten Skala des Drinking Refusal Self-Efficacy Questionnaire-Revised (DTSEQ-R; Oei et al., 2005) erfragt. Beurteilt wurde die Wahrscheinlichkeit, mit der die Probanden in 19 verschiedenen Situationen dem Konsum alkoholischer Getränke widerstehen können.

Psychische Gesundheit

Psychische Störungen

Psychischer Distress beschreibt eine latente Dimension, die bei Personen mit psychischen Störungen stark ausgeprägt ist und wurde mit einer von den Autoren ins Deutsche übersetzten Version des K6 (Kessler et al., 2002) erfasst. Das 6-Item Instrument erhebt, wie häufig Personen innerhalb der letzten 30 Tage Symptome von unspezifischem, sozioemotionalem Distress (z. B. Rastlosigkeit, Hoffungslosigkeit, Niedergeschlagenheit) erlebt haben und wird als Screening für psychische Störungen verwendet.

Manische Episode

Die Symptomatik einer manischen Episode wurde in Anlehnung an die Kriterien des DSM-IV (American Psychiatric Association, 1994) erhoben. In zwei Eingangskriterien wird zunächst erfasst, ob Personen innerhalb der letzten 12 Monate anhaltende Phasen von gehobener, expansiver oder

reizbarer Stimmung hatten. Werden beide Fragen positiv beantwortet, erfolgt in acht Items die Abfrage von manischen Symptomen (z. B. übersteigertes Selbstwertgefühl, vermehrte Gesprächigkeit, erhöhte Ablenkbarkeit).

Datenerhebung und -eingabe

Die Datenerhebung wurde zwischen Mai und Oktober 2009 durch das Institut für angewandte Sozialwissenschaften (infas) durchgeführt. Eingesetzt wurde ein Methodenmix aus schriftlicher, telefonischer und Internet-Befragung. Die Stichprobe wurde nach Verfügbarkeit der Telefonnummer der ausgewählten Adressen in einen schriftlichen und einen telefonischen Arm geteilt. Allen Zielpersonen des schriftlichen Arms wurde mit dem Anschreiben der Fragebogen postalisch zugesandt. Anschließend erfolgte ein gestuftes Verfahren, in dem alle Personen ohne Rücklauf dreimal postalisch an den Fragebogen erinnert wurden. Zusätzlich wurde den Zielpersonen über die Erinnerungsschreiben die Möglichkeit einer Beantwortung des Fragebogens am Telefon bzw. im Internet angeboten. Alle Nichtteilnehmer erhielten abschließend postalisch einen Nonresponse-Kurzfragebogen mit der Bitte, sechs Fragen zu ihrem Konsum von Tabak, Alkohol und illegalen Drogen sowie zu ihrem Spielverhalten schriftlich oder telefonisch zu beantworten.

Die Ermittlung der Telefonnummern erfolgte automatisiert (in aktuellen Telefonregistern) und für fehlende Einträge zusätzlich manuell (in älteren Telefonregistern). Alle Zielpersonen, für die eine Telefonnummer ermittelt werden konnte, erhielten vorab ein Anschreiben mit allen Informationen zur Studie, indem eine telefonische Kontaktaufnahme durch das Feldinstitut angekündigt wurde. Im Rahmen der telefonischen Kontaktierung, die bis zum Ende der Feldarbeiten lief, wurde Personen, die sich nicht sofort befragen ließen, die Möglichkeit einer Beantwortung des Fragebogens in schriftlicher Form bzw. im Internet angeboten. Jede Person, die erreicht wurde aber nicht teilnehmen wollte, wurde um die Beantwortung des Kurzfragebogens gebeten. An Zielpersonen, die nach 30 Kontaktversuchen nicht erreicht werden konnten oder die in der telefonischen Kontaktierung mitteilten, dass sie den Fragebogen nur schriftlich oder online beantworten wollten, erfolgte der postalische Versand des schriftlichen Fragebogens und des Online-Zugangscodes. Alle verbleibenden Nichtteilnehmer wurden zum Ende der Feldarbeiten gebeten, den Kurzfragebogen schriftlich zu beantworten.

Nach Rücklauf der Fragebögen (schriftlich bzw. online) und bei telefonisch durchgeführten Interviews erfolgte zunächst eine manuelle Prüfung hinsichtlich Vollständigkeit und Plausibilität. Die Datenerfassung des schriftlichen Fragebogens erfolgte anschließend maskengestützt durch professionelle CoderInnen mittels eines integrierten Computerprogramms, das Filtersprünge kontrolliert und eine Prüfung auf gültige Werte und Wertebereiche vornimmt. Bei telefonisch durchgeführten Interviews und im Internet be-

antworteten Fragebögen erfolgte die Datenerfassung direkt in die integrierte Software. Alle Daten wurden anschließend auf Filterfehler, Plausibilität und Inkonsistenzen überprüft und ggf. bereinigt.

Analysen

Auswertung

Alle Analysen wurden mit Stata 10.1 SE (Stata Corp LP, College Station, TX) durchgeführt, um die Besonderheiten des komplexen Stichprobendesigns zu berücksichtigen. Dies beinhaltet sowohl den Effekt des mehrstufigen Auswahlverfahrens (Clusterung) als auch die Gewichtung der Stichprobe. Zur Varianzschätzung wurde Taylorreihenentwicklung verwendet (Korn & Graubard, 1999).

Gewichtung

Die Gewichtung der Stichprobe ergibt sich durch Multiplikation zweier Gewichtungsvariablen (Korn & Graubard, 1999) und erlaubt die repräsentative Auswertung der Studienergebnisse hinsichtlich zentraler Merkmale in der Grundgesamtheit. Die designbedingte Disproportionalität unterschiedlicher Geburtsjahrganggruppen in der Stichprobe wurde durch ein Designgewicht ausgeglichen, das sich aus dem Kehrwert der Auswahlwahrscheinlichkeiten auf der jeweiligen Auswahlstufe ergibt. Zusätzlich wurde ein Poststratifikationsgewicht berechnet, das die gemeinsame Verteilung von Geburtsjahrgang und Geschlecht, sowie die Verteilungen von Bundesland und BIK-Gemeindegrößenklasse in der Grundgesamtheit anpasst. Der Mikrozensus 2008 bildete die Basis für die Daten zur Grundgesamtheit. Die Anpassung an Randverteilungen der Grundgesamtheit erfolgte über das Verfahren des Iterative Proportional Fitting (Gelman & Carlin, 2002).

Zur Beurteilung der Güte der Gewichtungen werden der Range und die Standardabweichung des Gewichtungsfaktors, das Effektivitätsmaß E und die effektive Stichprobengröße berichtet (Korn & Graubard, 1999; Little, Lewitzky, Heeringa, Lepkowski & Kessler, 1997). Letzteres ist definiert als $n' = (\sum g_i)^2 / \sum g_i^2$, wobei g = Gewichtungsvariable. E ist $n'/n \times 100$, wobei n = realisierte Fallzahl. Je näher E am Wert 100 % liegt, desto geringer ist die gewichtungsbedingte Reduktion der Fallzahl.

Berechnung der Antwortrate

Die Berechnung der Antwortrate erfolgte nach den Standards der American Association for Public Opinion Research (AAPOR, The American Association for Public Opinion Research, 2009). Dafür mussten die Anteile der nicht befragbaren Personen ermittelt werden. Grundlage hierfür bildeten alle Personen der Einsatzstichprobe, von

denen zum Ende der Feldarbeiten eine Information über die (Nicht-)Teilnahme an der Studie vorlag. Unterschieden wurde zwischen systematischen Ausfällen (d. h. Termin für telefonische Befragung nicht eingehalten, Zielperson verweigert die Teilnahme, ist während der Feldzeit nicht erreichbar oder aus gesundheitlichen Gründen nicht befragbar) und stichprobenneutralen Ausfällen (d. h. Telefonnummer unvollständig, Zielperson unbekannt verzogen, nicht deutschsprachig, verstorben oder nicht im Altersbereich der Befragung). Da der Anteil stichprobenneutraler Ausfälle an Personen, von denen zum Feldende keine Information vorlag, unbekannt ist, wurde eine Schätzung vorgenommen. Diese erfolgte auf Basis des Anteils stichprobenneutraler Ausfälle an den Befragten mit Informationen zum Feldende (Biemer & Lyberg, 2003). Die Bruttostichprobe wurde um den Anteil stichprobenneutraler Ausfälle (bekannt + geschätzt) bereinigt. Die Antwortrate ergibt sich dann als Verhältnis der Nettostichprobe zur bereinigten Bruttostichprobe.

Selektivitätsanalysen

Zur Abschätzung der Generalisierbarkeit der Studienergebnisse wurden zwei Formen von Selektivitätsanalysen durchgeführt. Im ersten Schritt wurden systematische Verzerrungen der realisierten Stichprobe aufgrund eines Zusammenhangs von Untersuchungsmerkmalen mit der Teilnahmebereitschaft untersucht. Hierzu wurde eine Nonresponse-Befragung zu wesentlichen Konsummerkmalen durchgeführt. Die Befragung erfolgte entweder in Form eines schriftlichen Kurzfragebogens (Nichtteilnehmer des schriftlichen Arms bzw. nicht erreichbare Zielpersonen des telefonischen Arms) oder direkt im Rahmen der telefonischen Kontaktierung für diejenigen Personen, die nicht zur Teilnahme an der Studie bereit waren. In Nonresponse-Analysen wurde geprüft, ob sich die Prävalenz des Rauchens und des Alkoholkonsums in den letzten 30 Tagen sowie die Lebenszeit- und die 12-Monats-Prävalenz des Cannabiskonsums zwischen Teilnehmern und Nichtteilnehmern, die den Nonresponse-Kurzfragebogen beantworteten, unterscheiden. Hierzu wurden logistische Regressionen gerechnet und die Konsumprävalenz jeweils als abhängige Variable modelliert, um Unterschiede zwischen Teilnehmern und Nichtteilnehmern, adjustiert nach Alter, Geschlecht, Bundesland und Interviewart (schriftlich, telefonisch, online) zu prüfen (Graubard & Korn, 1999).

Im zweiten Schritt der Selektivitätsanalyse wurden die realisierten Fälle hinsichtlich potentieller Modus-Effekte untersucht. Es wurde geprüft, ob in einzelnen Methoden bestimmte Konsumgruppen mit höherer Wahrscheinlichkeit vertreten sind als in anderen Methoden. Gegenübergestellt wurden die Angaben der telefonisch, schriftlich und im Internet befragten Zielpersonen zum Konsum von Alkohol, Tabak und Cannabis. Unterschiede in der Konsumprävalenz zwischen den einzelnen Methoden wurden mit logistischen Regressionen modelliert und nach Alter, Ge-

schlecht, Bundesland, Schulbildung und Haushaltsnettoeinkommen adjustiert.

Ergebnisse

Realisierte Stichprobe

Von 41.5 % ($n = 7.549$) der insgesamt gezogenen Einsatzadressen ($N = 18.176$) lag vor Beginn der Studie eine Telefonnummer vor. Die verbleibenden 58.5 % ($n = 10.627$) der Einsatzstichprobe erhielten den Fragebogen postalisch. Insgesamt nahmen 8.034 Personen (44.2 %) an der Befragung teil. Davon wurden 42.0 % ($n = 3.376$) der Interviews telefonisch realisiert, 46.4 % ($n = 3.731$) der Personen beantworteten den Fragebogen schriftlich und 11.5 % ($n = 927$) im Internet.

Bei Zielpersonen des schriftlichen Arms wurde mit dem Erstversand eine Realisierungsquote von 22.5 % ($n = 2.390$) erreicht. Die drei schriftlichen Erinnerungen erbrachten eine Steigerung von 15 Prozentpunkten, so dass insgesamt 38.8 % ($n = 4.122$) dieser Einsatzadressen realisiert werden konnten. Davon wurde mehr als ein Fünftel (21.9 %, $n = 901$) über den Internet-Fragebogen realisiert, der optional in jedem Erinnerungsschreiben angeboten wurde. Verglichen mit der Einsatzstichprobe des schriftlichen Arms war die Bruttorealisierung im telefonischen Arm mit 51.8 % ($n = 3.912$) deutlich höher. Davon wurden 86.3 % ($n = 3.376$) der Interviews telefonisch durchgeführt, 13.0 % ($n = 510$) der Zielpersonen beantworteten den auf Anfrage oder im Fall der telefonischen Nichterreichbarkeit übermittelten schriftlichen Fragebogen und 0.7 % ($n = 26$) den Fragebogen im Internet. Bis zum Feldende wurden durchschnittlich neun telefonische Kontaktversuche unternommen (minimal: 1; maximal: 158).

Stichprobenausfälle

Insgesamt 10.142 (55.8 %) Zielpersonen der Einsatzstichprobe ($N = 18.176$) nahmen nicht an der Studie teil. Von diesem Anteil konnten 14.8 % ($n = 1.502$) der Einsatzadressen als stichprobenneutrale Ausfälle identifiziert werden. Dieser Anteil gliederte sich in 950 (63.2 %) unbekannte Adressen, 435 (29.0 %) unvollständige Telefonnummern, 101 (6.7 %) nicht deutschsprachige und 5 (0.3 %) Zielpersonen außerhalb des Altersbereichs der Befragung sowie 11 (0.7 %) Verstorbene. Weitere 30.9 % ($n = 3.135$) der Zielpersonen wurden als systematische Ausfälle klassifiziert. Davon verweigerten 2.503 (79.8 %) Zielpersonen die Teilnahme schriftlich oder am Telefon, weitere 462 (14.7 %) Personen waren bis zum Ende der Feldzeit nicht erreichbar, aufgrund von Krankheit nicht befragbar ($n = 80$, 2.6 %) oder gaben an, den Fragebogen zu senden bzw. hielten den telefonischen Befragungstermin nicht ein ($n = 90$, 2.9 %). Von den verbleibenden 5.505 Zielpersonen

Tabelle 1

Demographische Verteilung der Bruttostichprobe, der Nettostichprobe (ungewichtet und gewichtet) und der Grundgesamtheit sowie die Ausschöpfungsraten

	Bruttostichprobe ohne neutrale Ausfälle		Ungewichtet			Gewichtet		Anteil an der Bevölkerung ¹
	n	%	n	Ausschöpfung	% Befragte	n	%	%
Insgesamt	16667	100.0	8030	49.8	100.0	8030	100.0	100.0
Männer, gesamt	8177	49.1	3558	45.2	44.3	4068	50.7	50.6
1989–1991	965	5.8	475	50.9	5.9	223	2.8	2.9
1985–1988	1152	6.9	447	40.7	5.6	310	3.9	3.8
1980–1984	1226	7.4	481	40.6	6.0	391	4.9	4.9
1970–1979	1573	9.4	648	42.7	8.1	791	9.9	10.2
1960–1969	1441	8.6	622	44.9	7.8	1111	13.8	13.8
1950–1959	1200	7.2	561	48.6	7.0	904	11.3	10.9
1945–1949	620	3.7	324	54.0	4.0	338	4.2	4.0
Frauen, gesamt	8490	50.9	4472	54.9	55.7	3962	49.3	49.4
1989–1991	978	5.9	584	61.2	7.3	212	2.6	2.8
1985–1988	1139	6.8	551	50.0	6.9	297	3.7	3.7
1980–1984	1280	7.7	627	50.2	7.8	381	4.8	4.8
1970–1979	1695	10.2	855	51.7	10.7	767	9.6	9.8
1960–1969	1443	8.7	772	54.6	9.6	1062	13.2	13.2
1950–1959	1244	7.5	700	58.0	8.7	898	11.2	10.9
1945–1949	701	4.2	383	56.9	4.8	345	4.3	4.2

¹Bevölkerung: Stand 31.12.2008.

(54.3 %) des schriftlichen Arms erfolgte trotz dreimaliger Erinnerung und dem zusätzlichen Angebot des Online-Fragebogens keine Reaktion.

Antwortrate

Für 69.7 % ($n = 12.671$) der Einsatzstichprobe ($N = 18.176$) lag nach Abschluss der Feldarbeiten eine Information über die (Nicht-) Teilnahme an der Studie vor. Der Anteil stichprobenneutraler Ausfälle an diesen Personen betrug 11.9 % ($n = 1.502$). Dieser Anteil diente als Schätzwert für die stichprobenneutralen Ausfälle an Personen, von denen bis zum Feldende keine Information vorlag ($n = 5.505$, 30.3 %). Aufgrund dieser Schätzung erhöhte sich die Anzahl stichprobenneutraler Ausfälle um $n = 653$ Personen, so dass insgesamt 2.155 Zielpersonen als neutrale Ausfälle angenommen wurden und sich die Bruttostichprobe auf 16.021 (88.1 %) Fälle reduzierte. Als Resultat der Datenprüfung wurden von 8.034 realisierten Fällen vier Personen ausgeschlossen, deren schriftliche Fragebögen nicht auswertbar waren. Somit verblieben 8.030 Fragebögen/Interviews als Basis für die Datenauswertung. Nach Abzug der stichprobenneutralen Ausfälle ergibt sich somit eine Antwortrate von 50.1 %. Werden die neutralen Ausfälle bei der Berechnung nicht berücksichtigt, beträgt die Antwortrate 44.2 %.

Güte der Gewichtung

Die verwendete Gewichtungsvariable, bestehend aus Design- und Poststratifikationsfaktor, hat einen Range von 0.27 bis 2.90 und eine Standardabweichung von 0.44. Das Designgewicht allein weist dabei ein Minimum von 0.41 und ein Maximum von 1.56 sowie eine Standardabweichung von 0.40 auf. Das Effektivitätsmaß der verwendeten Gewichtung beträgt 83.5 %, was einer effektiven Stichprobengröße von 6.704 Fällen entspricht. Die demographische Verteilung der Bruttostichprobe, der ungewichteten und gewichteten Nettostichprobe sowie der Grundgesamtheit ist gegliedert nach Jahrganggruppen in Tabelle 1 dargestellt. Aufgrund des disproportionalen Stichprobendesigns sind in der ungewichteten Stichprobe bei beiden Geschlechtern jüngere Personen über- und ältere Personen im Vergleich zur Grundgesamtheit unterrepräsentiert. Weiterhin ist der Anteil weiblicher Befragter in allen Altersgruppen deutlich höher als der Anteil männlicher Befragter. Die Ausschöpfungsraten sind insbesondere bei jüngeren und älteren Befragten beiden Geschlechts höher als im mittleren Erwachsenenalter. In der gewichteten Stichprobe stimmen die Altersgruppenverteilungen bei Männern und Frauen sehr gut mit der Verteilung in der Grundgesamtheit überein.

Die Verteilungen sozioökonomischer Merkmale in der ungewichteten und gewichteten Stichprobe sind in Tabelle 2 denen in der Grundgesamtheit gegenübergestellt. Aufgrund des disproportionalen Designs ist der Anteil lediger

Tabelle 2

Verteilung soziodemographischer Merkmale in der ungewichteten und gewichteten Stichprobe und in der Grundgesamtheit

		Ungewichtete Stichprobe			Gewichtete Stichprobe			Bevölkerung ¹ (18- bis 64-Jährige)		
		Gesamt	Männer	Frauen	Gesamt	Männer	Frauen	Gesamt	Männer	Frauen
Staatsbürgerschaft	deutsch	93.7	93.6	93.8	94.1	94.0	94.1	89.4	89.3	89.4
	andere	6.2	6.3	6.1	5.8	5.9	5.8	10.6	10.7	10.6
	keine Angabe	0.1	0.1	0.1	0.1	0.1	0.1			
Familienstand	ledig	46.6	49.6	44.2	34.3	37.2	31.2	37.8	42.3	33.3
	verheiratet	45.0	43.6	46.1	55.5	54.4	56.5	50.3	47.6	53.0
	verwitwet	1.5	0.7	2.1	1.6	0.8	1.6	2.0	0.9	3.2
	geschieden	6.0	5.2	6.6	7.8	6.8	7.8	9.9	9.2	10.5
	keine Angabe	0.9	0.9	1.0	0.9	0.8	0.9			
Schulabschluss	in Schulausbildung	4.2	3.9	4.5	2.3	2.1	2.4	1.6	1.6	1.6
	Hauptschulabschluss	15.6	18.1	13.7	17.7	19.8	15.5	31.0	33.2	28.8
	Polytechnische Oberschule	4.9	4.6	5.1	6.3	6.0	6.8	8.7	8.5	8.9
	Mittlere Reife	28.6	25.8	30.8	27.4	24.0	30.8	24.6	21.7	27.6
	(Fach-)Abitur	43.6	44.7	42.7	43.0	45.0	40.8	29.6	30.7	28.5
	ohne Abschluss	1.3	1.5	1.2	1.5	1.6	1.3	3.7	3.6	3.8
	anderer Abschluss	1.3	1.0	1.6	1.4	1.2	1.7	0.4	0.4	0.4
	keine Angabe	0.5	0.4	0.6	0.5	0.4	0.6	0.3	0.3	0.3
Haushaltsnettoeinkommen ²	unter 500	2.7	3.0	2.6	2.2	2.4	1.9			
	500–1.000	8.1	7.3	8.7	7.1	6.6	7.5			
	1.000–1.250	5.5	4.4	6.3	5.0	4.2	5.8			
	1.250–1.500	7.1	6.4	7.6	6.8	6.1	7.6			
	1.500–2.000	13.8	13.6	13.9	13.6	13.2	13.9			
	2.000–2.500	14.7	14.5	14.9	15.5	15.3	15.6			
	2.500–3.000	12.0	11.5	12.3	13.4	11.8	12.9			
	3.000–4.000	14.0	15.4	12.8	15.3	16.8	13.7			
	4.000 und mehr	15.4	17.4	13.9	17.1	18.5	15.6			
keine Angabe	6.8	6.5	7.0	5.2	4.9	5.5				

¹Bevölkerung: Stand 31.12.2008. ²im Statistischen Jahrbuch abweichende Einkommensgrenzen, daher kein Vergleich mit der Grundgesamtheit möglich.

und in Schulausbildung befindlicher Personen in der ungewichteten Stichprobe im Vergleich zur Grundgesamtheit überrepräsentiert, stimmt jedoch in der gewichteten Stichprobe besser mit der Verteilung in der Bevölkerung überein. Sowohl in der gewichteten als auch in der ungewichteten Stichprobe ist der Anteil an Personen mit Hauptschulabschluss deutlich geringer und der Anteil an Personen mit Abitur deutlich größer als in der Grundgesamtheit. Unterrepräsentiert sind zudem nicht-deutsche Staatsangehörige. Aufgrund abweichender Einkommensgrenzen im Mikrozensus ist kein Vergleich der ESA-Stichprobe mit der Grundgesamtheit hinsichtlich der Angaben zum Haushaltsnettoeinkommen möglich.

Selektivitätsanalysen

Von den 10.142 Personen, die nicht an der Studie teilnahmen, wurden alle potentiell befragbaren Nichtteilnehmer

($n = 8.927$, 88.0 %) gebeten, sechs kurze Fragen zum Konsum von Alkohol, Tabak und illegalen Drogen sowie zum Glücksspielverhalten zu beantworten. Nicht befragbar und damit von der Nonresponse-Befragung ausgeschlossen waren Personen, die entweder unbekannt verzogen ($n = 950$), verstorben ($n = 11$), nicht im angestrebten Altersbereich ($n = 5$), nicht deutschsprachig ($n = 90$) oder aus gesundheitlichen Gründen nicht befragbar waren ($n = 69$) sowie den telefonischen Interviewtermin nicht einhielten ($n = 39$) oder den Fragebogen schriftlich senden wollten ($n = 51$). Von allen Personen, welche die Nonresponse-Fragen erhielten, beantworteten insgesamt 704 Personen (7.9 %) die sechs Fragen, davon 10.8 % ($n = 76$) telefonisch und 89.2 % ($n = 628$) schriftlich.

Im Rahmen der Nonresponse-Analyse wurden ausgewählte Konsumvariablen zwischen Teilnehmern und Nichtteilnehmern verglichen (Tabelle 3). In Bezug auf Alkohol- und Drogenkonsum zeigten sich Unterschiede in Abhängigkeit von der Teilnahmebereitschaft. Der Anteil

Tabelle 3
Adjustierte¹ Konsumvariablen nach Teilnahmebereitschaft (in %)

	Teilnehmer <i>n</i> = 8030	Nichtteilnehmer <i>n</i> = 704
Tabakkonsum, letzte 30 Tage	28.9	31.6
Alkoholkonsum, letzte 30 Tage	76.7	72.9*
Cannabiskonsum, Lebenszeit	24.6	10.9*
Cannabiskonsum, letzte 12 Monate	3.4	1.6*

¹Die Adjustierung nach Alter, Geschlecht, Bundesland, und Interviewart wurde mit logistischen Regressionsmodellen vorgenommen. **p* < .05 für Vergleich mit «Teilnehmer».

Tabelle 4
Adjustierte¹ Konsumvariablen nach Art der Befragung (in %)

	Schriftlich <i>n</i> = 3731	Telefon <i>n</i> = 3376	Internet <i>n</i> = 927
Tabakkonsum, letzte 30 Tage	28.7	28.5	28.7
Alkoholkonsum, letzte 30 Tage	78.9	77.4	74.7*
Cannabiskonsum, Lebenszeit	26.9	20.8*	25.4
Cannabiskonsum, letzte 12 Monate	3.3	2.4*	3.5*

¹Die Adjustierung nach Alter, Geschlecht, Bundesland, Schulbildung und Haushaltsnettoeinkommen wurde mit logistischen Regressionsmodellen vorgenommen.

**p* < .05 für Vergleich mit «Schriftlich».

der Nichtteilnehmer, die in den letzten 30 Tagen Alkohol konsumiert hatten, ist im Vergleich zu den Teilnehmern an der Studie signifikant geringer. Ebenso berichteten Nichtteilnehmer deutlich seltener, jemals im Leben sowie in den letzten 12 Monaten Cannabis konsumiert zu haben. Vergleiche zwischen den Konsumangaben der Teilnehmer und Nichtteilnehmer zeigten keine signifikanten Unterschiede hinsichtlich der Prävalenz des Tabakkonsums in den letzten 30 Tagen.

Die Konsumangaben der realisierten Fälle unterscheiden sich zum Teil zwischen den einzelnen Modi der Befragung (Tabelle 4). Verglichen mit den schriftlich befragten Teilnehmern weisen telefonisch befragte Personen eine geringere Lebenszeit- und 12-Monats-Prävalenz des Cannabiskonsums auf. Umgekehrt ist der Anteil an Cannabiskonsumern der letzten 12 Monate in der Internet-Befragung höher und der Anteil an Alkoholkonsumenten der letzten 30 Tage geringer im Vergleich zur postalischen Befragung.

Diskussion

Die erzielte Ausschöpfung von 50.1 % nach Abzug der stichprobenneutralen Ausfälle kann, gemessen an der generell verminderten Teilnahmebereitschaft an Bevölke-

rungsumfragen in den letzten Jahren (De Leeuw & de Heer, 2002; Schnell, 1997), als befriedigend gewertet werden. Reine Telefonbefragungen wie der bislang letzte bundesweite Telefongesundheitssurvey erreichen höhere Ausschöpfungsquoten (56 %; Robert Koch-Institut, 2006), mündliche Interviews wie die letzte Erhebung des ALLBUS von 2008 hingegen niedrigere Antwortraten (40 %; Wasmer, Scholz & Blohm, 2010). Als Gründe für die gestiegene Unit-Nonresponse in Bevölkerungsumfragen werden unter anderem fehlende Motivation, Zeitmangel und Bedenken in Bezug auf die Privatsphäre berichtet (Schnell, 1997). Zudem diskutiert Schnell (1997) Unterschiede in der Erreichbarkeit als Hauptursache für die im ESA gefundenen demographischen Unterschiede (Geschlecht, Alter) in der Nonresponse. Die geringere Ausschöpfung unter Männern und Erwachsenen mittleren Alters ist demnach nicht notwendigerweise Ergebnis einer geringeren Teilnahmebereitschaft unter diesen Personen. Als Motiv für die Nichtteilnahme ist aber auch sog. «Oversurveying» d. h. die Überladung mit Teilnahmeanfragen denkbar (Schnauber & Daschmann, 2008). Wichtig wird es daher sein, auch bei zukünftigen Erhebungen deutlich den wissenschaftlichen Charakter des Epidemiologischen Suchtsurveys zu betonen und die Studie klar von kommerzieller Forschung und Meinungsumfragen abzugrenzen.

Positiv konstatiert werden kann der Anstieg der Antwortrate um fünf Prozentpunkte gegenüber dem ESA 2006, was auf den in der vorliegenden Studie erstmals von Beginn an eingesetzten Methodenmix aus telefonischer und schriftlicher Befragung zurückzuführen sein dürfte. Generell erzielten Telefoninterviews höhere Ausschöpfungen als schriftliche Befragungen (Hox & de Leeuw, 1994) und zeichnen sich durch geringere Missingraten aus (Brogger, Bakke, Eide & Gulsvik, 2002). So ist der Realisierungsanteil an Einsatzadressen mit Telefonnummer im ESA 2009 erwartungsgemäß deutlich größer als der Realisierungsanteil an schriftlich Befragten. Auch das zusätzliche Angebot der Online-Beantwortung des Fragebogens zeigte einen deutlichen positiven Effekt auf die Teilnahmebereitschaft. In Folgestudien sollten dennoch weitere Maßnahmen ergriffen werden, um die Antwortrate zu steigern und Verzerrungen durch die Teilnahmebereitschaft zu minimieren. Vor allem ein kürzerer Fragebogen (Nakash, Hutton, Jorstad-Stein, Gates & Lamb, 2006) und telefonische statt schriftliche Erinnerungen (Salim, Smith & Bammer, 2002) sollten dabei erwogen werden. Auch der Einsatz von Anreizen ist erwiesenermaßen geeignet, die Antwortrate zu steigern (Edwards et al., 2007). So berichten Jobber, Saunders und Mitchell (2004), dass sich die Rücklaufquote im Schnitt um 15 Prozentpunkte durch Einsatz eines Anreizes erhöht, unabhängig von dessen materiellem Wert.

Eine Gewichtung der Daten des Epidemiologischen Suchtsurveys ist generell erforderlich, um Abweichungen in den Verteilungen zentraler Merkmale von der Bevölkerung auszugleichen. Die hier verwendete Gewichtung erzielt mit einer Effektivität von 84 % eine gute Anpassung (Korn & Graubard, 1999) und ist mit den Gesundheitssur-

veys des Robert Koch-Instituts vergleichbar (Kohler, 2005). Abweichungen von der Grundgesamtheit bestehen allerdings auch in der gewichteten Stichprobe hinsichtlich Staatsbürgerschaft und Schulbildung. Da der Fragebogen auf Deutsch konzipiert wurde und nicht deutschsprachige Zielpersonen von der Befragung ausgeschlossen wurden, sind Personen anderer Herkunft im Survey unterrepräsentiert. Weiterhin ist das Bildungsniveau unter den Teilnehmern des ESA 2009 im Vergleich zur Allgemeinbevölkerung deutlich höher. Dieser «Mittelschichtbias» ist unter anderem auf eine eher misstrauische Einstellung gegenüber Umfragen sowie ein geringeres thematisches Interesse bei weniger gebildeten Personen zurückzuführen (Diekmann, 2007; Schnell, 1997).

Mit einer Antwortrate von 50.1 % besteht die Möglichkeit, dass die ausgewerteten Fälle nicht mehr repräsentativ für die Grundgesamtheit sind, da die Populationsschätzer aufgrund von Selektionseffekten verfälscht werden. Im Rahmen einer Nonresponse-Analyse wurde daher geprüft, inwieweit sich Nichtteilnehmer von Teilnehmern der Studie bezüglich relevanter Konsummerkmale unterscheiden. Es zeigte sich, dass Nichtteilnehmer seltener Alkohol- und Drogenkonsum berichten als Teilnehmer. Damit sind aktuelle Alkoholkonsumenten und Personen mit Erfahrungen hinsichtlich des Konsums von Cannabis eher in der Studie repräsentiert. Im Einklang mit diesen Ergebnissen zeigten auch andere Studien (Lahaut et al., 2003), dass Verweigerer eher alkoholabstinent sind als Studienteilnehmer. Nicht ausschließen lässt sich jedoch, dass Alkoholkonsumenten, die die Teilnahme an der Studie verweigerten, eher episodisch große Alkoholmengen konsumieren (Lemmens, Tan & Knibbe, 1988) und Rauschtrinken im ESA 2009 damit unterschätzt wird. In Bezug auf den Konsum von Cannabis legen die Analysen von Zhao, Stockwell und Macdonald (2009) hingegen höhere Prävalenzwerte bei Nichtteilnehmern nahe. Im Gegensatz dazu weisen die Befunde des ESA 2009 eher auf eine höhere Verleugnungstendenz als Erklärung der geringeren Prävalenzen des Cannabiskonsums unter Verweigerern hin. Einschränkend ist jedoch anzumerken, dass nur 8 % der Verweigerer den Kurzfragebogen beantworteten, wodurch weitere Selektionseffekte nicht auszuschließen sind.

Weiterhin können Verzerrungen im Antwortverhalten durch Unterschiede zwischen den einzelnen Befragungsmodi bedingt sein. Die Ergebnisse der Selektivitätsanalyse zeigen, dass Cannabiskonsum unter telefonisch Befragten seltener verbreitet war als unter schriftlich Befragten. Dieser Befund wurde auch in anderen Untersuchungen berichtet (Aquilino, 1994; Aquilino & LoSciuto, 1990) und deutet auf eine subjektiv geringere Anonymität bei telefonischen Interviews und damit eine höhere Verleugnungstendenz hin. Der Vergleich zwischen schriftlich und online Befragten zeigt, dass letztere seltener Alkoholkonsumenten waren, jedoch häufiger Cannabiskonsum in den letzten 12 Monaten angaben. Dieser Befund war unter anderem unabhängig von Alter, Geschlecht und Schulbildung der Befragten zu beobachten. Allerdings können Verzerrungen

aufgrund unterschiedlicher Zugangsmöglichkeiten zum Internet nicht ausgeschlossen werden (Bandilla, Kaczmirek, Blohm & Neubarth, 2009).

Der Epidemiologische Suchtsurvey 2009 stellt trotz Limitationen ein wichtiges Instrument zur Beschreibung von Umfang und Häufigkeit des Konsums von legalen und illegalen Substanzen in der Bevölkerung dar. Neben den seit vielen Jahren erhobenen epidemiologischen Basisdaten zu Konsum, Gebrauchsmustern und substanzbezogenen Störungen liegen mit dem ESA 2009 auch aktuelle Zahlen zum Glückspielverhalten und zum Ausmaß pathologischen Spielverhaltens in der Bevölkerung vor. Weitere Anstrengungen müssen jedoch unternommen werden, um die Antwortrate in zukünftigen Erhebungen weiter zu steigern.

Danksagung

Der Epidemiologische Suchtsurvey (ESA) wurde aus Mitteln des Bundesministeriums für Gesundheit (BMG) gefördert (Förderkennzeichen: 119–4914–8/32). Mit der Finanzierung sind keine Auflagen verbunden.

Deklaration möglicher Interessenkonflikte

Es liegen keinerlei Interessenkonflikte im Zusammenhang mit dieser Publikation vor.

Literatur

- American Psychiatric Association. (1994). *DSM-IV Diagnostic and statistical manual of mental disorders*. Washington, DC: Autor.
- Aquilino, W. S. (1994). Interview mode effects in surveys of drug and alcohol use: A field experiment. *Public Opinion Quarterly*, 58, 210–240.
- Aquilino, W. S. & LoSciuto, L. A. (1990). Effects of interview mode on self-reported drug use. *Public Opinion Quarterly*, 54, 362–395.
- Babor, T. F., Higgins-Biddle, J. C., Saunders, J. B. & Monteiro, M. G. (2001). *AUDIT. The Alcohol Use Disorders Identification Test. Guidelines for use in primary care*. Geneva: World Health Organisation.
- Bandilla, W., Kaczmirek, L., Blohm, M. & Neubarth, W. (2009). Coverage- und Nonresponse-Effekte bei Online-Bevölkerungsumfragen. In N. Jakob, H. Schoen & T. Zerback (Hrsg.), *Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung* (S. 129–143). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.
- Baumeister, S., Kraus, L., Stonner, T. & Metz, K. (2008). Tabakkonsum, Nikotinabhängigkeit und Trends. Ergebnisse des Epidemiologischen Suchtsurveys 2006. *Sucht*, 54(Sonderheft 1), S26–S35.

- Biemer, P. & Lyberg, L. (2003). *Introduction to survey quality*. New York: Wiley and Sons.
- Brogger, J., Bakke, P., Eide, G. E. & Gulsvik, A. (2002). Comparison of telephone and postal survey modes on respiratory symptoms and risk factors. *American Journal of Epidemiology*, 155, 572–576.
- Bühlinger, G., Augustin, R., Bergmann, E., Bloomfield, K., Funk, W., Junge, B., . . . Töppich, J. (2000). *Alkoholkonsum und alkoholbezogene Störungen in Deutschland* (Schriftenreihe des Bundesministeriums für Gesundheit Bd. 128). Baden-Baden: Nomos.
- Bundeszentrale für gesundheitliche Aufklärung. (2007). *Alkoholkonsum der Jugendlichen in Deutschland 2004 bis 2007*. Köln: Bundeszentrale für gesundheitliche Aufklärung.
- Bundeszentrale für gesundheitliche Aufklärung. (2008). *Die Drogenaffinität Jugendlicher in der Bundesrepublik Deutschland 2008. Alkohol-, Tabak- und Cannabiskonsum. Erste Ergebnisse zu aktuellen Entwicklungen und Trends*. Köln: Bundeszentrale für gesundheitliche Aufklärung.
- Currie, S. R., Hodgins, D. C., Wang, J., El-Guebaly, N., Wynne, H. & Chen, S. (2006). Risk of harm among gamblers in the general population as a function of level of participation in gambling activities. *Addiction*, 101, 570–580.
- De Leeuw, E. & de Heer, W. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge & J. A. Roderick (Hrsg.), *Survey nonresponse*. New York: Wiley and Sons.
- Diekman, A. (2007). *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen* (17. Auflage). Reinbeck: Rowohlt.
- Edwards, P., Roberts, I., Clarke, M., DiGiuseppi, C., Pratap, S., Wentz, R., . . . Cooper, R. (2007). Methods to increase response rates to postal questionnaires. *Cochrane Database Systematic Reviews* 2007, Issue 2. Art. No.: MR000008. DOI: 10.1002/14651858.MR000008.pub3.
- Etter, J.-F. & Perneger, T. V. (1999). A comparison of two measures of stage of change for smoking cessation. *Addiction*, 94, 1881–1889.
- Franke, A., Mohn, K., Sitzler, F., Welbrink, A. & Witte, M. (2001). *Alkohol- und Medikamentenabhängigkeit bei Frauen. Risiken und Widerstandsfaktoren*. Weinheim: Juventa.
- Gelman, A. & Carlin, J. (2002). Poststratification and weighting adjustments. In R. M. Groves, J. L. Eltinge & R. J. A. Little (Hrsg.), *Survey nonresponse* (S. 289–203). New York: Wiley.
- Gmel, G. & Rehm, J. (2004). Measuring alcohol consumption. *Contemporary Drug Problems*, 34, 467–540.
- Gmel, G., Rehm, J. & Kuntsche, E. (2003). Binge drinking in Europe: Definitions, epidemiology, and consequences. *Sucht*, 49, 105–116.
- Gossop, M., Darke, S., Griffiths, P., Hando, J., Powis, B., Hall, W. & Strang, J. (1995). The Severity of Dependence Scale (SDS): Psychometric properties of the SDS in English and Australian samples of heroin, cocaine and amphetamine users. *Addiction*, 90, 607–614.
- Graubard, B. I. & Korn, E. L. (1999). Predictive margins with survey data. *Biometrics*, 55, 652–659.
- Hagen, J. & Demmel, R. (2002). Inventar akuter Alkoholeffekte (IAA). In A. Glöckner-Rist, F. Rist & H. Kufner (Hrsg.), *Elektronisches Handbuch zu Erhebungsinstrumenten im Suchtbereich (EHES) (Version 2.00)*. Mannheim: Zentrum für Umfragen, Methoden und Analysen e. V.
- Hagen, J., Demmel, R. & Rist, F. (2002). Das Inventar akuter Alkoholeffekte: Faktorenstruktur und psychometrische Eigenschaften. In G. Richter, H. Rommelspacher & C. Spies (Hrsg.), *«Alkohol, Nikotin, Kokain . . . und kein Ende?»* (S. 543). Lengerich: Pabst Science Publisher.
- Heatherston, T. F., Kozlowski, L. T., Frecker, R. C. & Fagerström, K. O. (1991). The Fagerström test for nicotine dependence: A revision of the Fagerström tolerance questionnaire. *British Journal of Addiction*, 86, 1119–1127.
- Hox, J. J. & de Leeuw, E. D. (1994). A comparison of nonresponse in mail, telephone, and face-to-face surveys. *Quality and Quantity*, 28, 329–344.
- Jobber, D., Saunders, J. & Mitchell, V. W. (2004). Prepaid monetary incentive effects on mail survey response. *Journal of Business Research*, 57, 21–25.
- Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S. L., . . . Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, 32, 959–976.
- Kohler, M. (2005). Stichprobenkonzeption und Stichprobenrealisierung. In M. Kohler, A. Rieck, S. Borch & T. Ziese (Hrsg.), *Erster telefonischer Gesundheitssurvey des Robert-Koch-Instituts – Methodische Beiträge* (S. 31–57). Berlin: Robert-Koch-Institut.
- Korn, E. L. & Graubard, B. I. (1999). *Analysis of health surveys*. New Jersey: Wiley and Sons.
- Kraus, L. & Baumeister, S. (2008). Studiendesign und Methodik des Epidemiologischen Suchtsurveys 2006. *Sucht*, 54(Sonderheft 1), S6–S15.
- Kraus, L., Baumeister, S. E., Pabst, A. & Orth, B. (2009). Association of average daily alcohol consumption, binge drinking and alcohol-related social problems: Results from the German Epidemiological Surveys of Substance Abuse. *Alcohol and Alcoholism*, 44, 314–320.
- Lachner, G., Wittchen, H. U., Perkonig, A., Holly, A., Schuster, P., Wunderlich, U., . . . Pfister, H. (1998). Structure, content and reliability of the Munich-Composite International Diagnostic Interview (M-CIDI) substance use sections. *European Addiction Research*, 4, 28–41.
- Lahaut, V. M., Jansen, H. A., van de Mheen, D., Garretsen, H. F., Verdurmen, J. E. & van Dijk, A. (2003). Estimating non-response bias in a survey on alcohol consumption: Comparison of response waves. *Alcohol and Alcoholism*, 38, 128–134.
- Lemmens, P. H., Tan, E. S. & Knibbe, R. A. (1988). Bias due to non-response in a Dutch survey on alcohol consumption. *British Journal of Addiction*, 83, 1069–1077.
- Little, R. J. A., Lewitzky, S., Heeringa, S., Lepkowski, J. & Kessler, R. C. (1997). Assessment of weighting methodology for the National Comorbidity Survey. *American Journal of Epidemiology*, 146, 439–449.
- Müller, S., Piontek, D., Pabst, A., Baumeister, S. E. & Kraus, L. (2010). Changes in alcohol consumption and beverage preference among adolescents after the introduction of the alcopops tax in Germany. *Addiction*, 105, 1205–1213.
- Nakash, R. A., Hutton, J. L., Jorstad-Stein, E. C., Gates, S. & Lamb, S. E. (2006). Maximising response to postal questionnaires—a systematic review of randomised trials in health research. *Medical Research Methodology*, 6, 5. DOI: 10.1186/1471-2288-6-5.
- Oei, T. P., Hasking, P. A. & Young, R. M. (2005). Drinking refusal self-efficacy questionnaire-revised (DRSEQ-R): A new factor

- structure with confirmatory factor analysis. *Drug and Alcohol Dependence*, 78, 297–307.
- Pabst, A., Baumeister, S. E. & Kraus, L. (2010). Alcohol-expectancy dimensions and alcohol consumption at different ages in the general population. *Journal of Studies on Alcohol and Drugs*, 71, 46–53.
- Rist, F., Scheuren, B., Demmel, R., Hagen, J. & Aulhorn, I. (2003). Der Münsteraner Alcohol Use Disorders Identification Test (AUDIT-G-M). In A. Glöckner-Rist, F. Rist & H. Kufner (Hrsg.), *Elektronisches Handbuch zu Erhebungsinstrumenten im Suchtbereich (EHES). Version 3.00*. Mannheim: Zentrum für Umfragen, Methoden und Analysen.
- Robert Koch-Institut. (2006). *Bundesweiter Telefongesundheitsurvey (4. Welle – GSTel06)*. Verfügbar unter http://www.rki.de/cln_169/nn_201180/DE/Content/GBE/Erhebungen/Gesundheitsurvey/Geda/Cati06_inhalt.html
- Salim, S. M., Smith, W. T. & Bammer, G. (2002). Telephone reminders are a cost effective way to improve responses in postal health surveys. *Journal of Epidemiology and Community Health*, 56, 115–118.
- Schär, M., Hornung, R., Gutscher, H. & May, U. (2007). *Selbstmedikation – Ergebnisse einer Repräsentativerhebung im Rahmen des Nationalen Forschungsprogramms Nr. 8*. Zürich: Institut für Sozial- und Präventivmedizin der Universität Zürich.
- Schnauber, A. & Daschmann, G. (2008). States oder Traits? Was beeinflusst die Teilnahmebereitschaft an telefonischen Interviews? *Methoden – Daten – Analysen*, 2, 97–123.
- Schnell, R. (1997). *Nonresponse in Bevölkerungsumfragen, Ausmaß, Entwicklung und Ursachen*. Opladen: Leske + Budrich.
- Statistisches Bundesamt. (2004). *Demographische Standards*. Wiesbaden: Statistisches Bundesamt.
- Steiner, S., Baumeister, S. & Kraus, L. (2008). Severity of Dependence Scale – Establishing a cut-off point for cannabis dependence in the German adult population. *Sucht*, 54(Sonderheft 1), S57–S63.
- Stinchfield, R. (2002). Reliability, validity, and classification accuracy of the South Oaks Gambling Screen (SOGS). *Addictive Behaviors*, 27, 1–19.
- The American Association for Public Opinion Research. (2009). *Standard definitions: Final dispositions of case codes and outcome rates for surveys* (6th edition). Lenexa, KS: AAPOR.
- Wasmer, M., Scholz, E. & Blohm, M. (2010). *Konzeption und Durchführung der «Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften» (ALLBUS) 2008*. GESIS Technical Report 2010/04. Mannheim: GESIS – Leibniz-Institut für Sozialwissenschaften.
- Watzl, H., Rist, F., Höcker, W. & Miehle, K. (1991). Entwicklung eines Fragebogens zur Erfassung von Medikamentenmißbrauch bei Suchtpatienten. In M. Heide & H. Lieb (Hrsg.), *Sucht und Psychosomatik. Beiträge des 3. Heidelberger Kongresses* (S. 123–139). Bonn: Nagel.
- Wittchen, H.-U., Beloch, E., Garczynski, E., Holly, A., Lachner, G., Perkonig, A., ... Ziegler, S. (1995). *Münchener Composite International Diagnostic Interview (M-CIDI), Paper-pencil 2.2, 2/95*. München: Max-Planck-Institut für Psychiatrie, Klinisches Institut.
- Zhao, J., Stockwell, T. & Macdonald, S. (2009). Nonresponse bias in alcohol and drug population surveys. *Drug and Alcohol Review*, 28, 648–657.

Eingereicht: 11.06.2010

Angenommen: 17.09.2010

PD Dr. Ludwig Kraus

IFT Institut für Therapieforchung
 Leiter des Fachbereichs Epidemiologische Forschung
 Parzivalstr. 25
 D-80804 München
 Tel. +49 89 3-608-0430
 Fax +49 89 3-608-0449
 E-Mail: kraus@ift.de