

Themenschwerpunkt

Studiendesign und Methodik des Epidemiologischen Suchtsurveys 2012

Ludwig Kraus^{1,2}, Daniela Piontek¹, Alexander Pabst¹ und Elena Gomes de Matos¹¹IFT Institut für Therapieforschung, München²Centre for Social Research on Alcohol and Drugs (SoRAD), Stockholm University, Stockholm

Zusammenfassung: *Ziel:* Es wird ein Überblick über Design und Stichprobenziehung, Erhebungsinstrumente, Durchführung sowie methodische Analysen des Epidemiologischen Suchtsurveys (ESA) 2012 gegeben. *Methodik:* Die Ziehung der Personenstichprobe erfolgte disproportional zur Verteilung der Geburtsjahrgänge auf Basis der Einwohnermelderegister in einem zweistufigen Zufallsverfahren. Die Befragungen wurden im Methodenmix durchgeführt (schriftlich, telefonisch, Internet). *Ergebnisse:* Insgesamt nahmen 9084 Personen im Alter zwischen 18 und 64 Jahren an der Studie teil (Nettoausschöpfung 53.6 %). Das Redressement-Gewicht hatte eine Effektivität von 67.6 %. Nicht-Teilnehmer wiesen eine niedrigere Konsumprävalenz von Alkohol und Cannabis, aber eine erhöhte Prävalenz des episodischen Rauschtrinkens und Schlafmittelgebrauchs sowie höhere Tabakkonsummengen auf als Teilnehmer. Telefonisch und im Internet befragte Personen gaben seltener Substanzkonsum an als schriftlich Befragte. *Schlussfolgerungen:* Der zentralen Herausforderung sinkender Ausschöpfungsraten in epidemiologischen Studien kann der Survey durch den Einsatz unterschiedlicher Befragungsmethoden standhalten. Selektionseffekte sind aufgrund der Nicht-Berücksichtigung bestimmter Bevölkerungsgruppen sowie aufgrund von Non-response-Effekten möglich.

Schlüsselwörter: Bevölkerungsbefragung, Methodik, Substanzkonsum, substanzbezogene Störungen, Non-response

Study Design and Methodology of the 2012 Epidemiological Survey of Substance Abuse

Abstract: *Aims:* The paper gives an overview on design and sample selection, measures, realization and analyses of the 2012 Epidemiological Survey of Substance Abuse (ESA). *Methods:* A disproportional sample was drawn from population registers using a two-stage probability design oversampling younger birth cohorts. Different modes of administration were used (paper-and-pencil questionnaire, telephone interview, online questionnaire). *Results:* A total of 9 084 individuals aged 18 to 64 years participated in the survey (response rate 53.6 %). The redressement weight had an effectiveness of 67.6 %. Non-responders showed a lower prevalence of alcohol and cannabis use, but a higher prevalence of episodic heavy drinking and hypnotics use as well as a higher tobacco use quantity. Respondents in the telephone and internet mode had a lower rate of substance use than those in the paper-and-pencil mode. *Conclusions:* Contrary to the general trend of decreasing response rates, the ESA achieves increasing response rates over the past years due to the application of a mixed mode design. Selectivity effects are possible because of the exclusion of specific population subgroups and non-response effects.

Keywords: General population survey, methodology, substance use, substance use disorders, non-response

Einführung

Der Epidemiologische Suchtsurvey (ESA) wurde im Jahr 2012 zum zehnten Mal durchgeführt und umfasst seit der ersten Erhebung im Jahr 1980 nunmehr eine Zeitspanne von mehr als 30 Jahren. Dabei hat der Survey über die Zeit unterschiedliche methodische Anpassungen erfahren, um sich ändernden Anforderungen zu entsprechen. Ursprünglich als Studie für das Jugend- und junge Erwachsenenalter (12 bis 24 Jahre) durchgeführt, wurde das untere

Alterslimit ab 1995 auf 18 Jahre erhöht und das obere Alterslimit in mehreren Stufen auf 29 Jahre (1986), 39 Jahre (1990–1992), 59 Jahre (1995–2003) und schließlich 64 Jahre (seit 2006) angehoben. Seit 1990 wird der Survey in West- und Ostdeutschland durchgeführt. Um die zunehmend schiefe Altersverteilung in der Bevölkerung auszugleichen, wird seit dem Jahr 2003 eine disproportional Stichprobe gezogen, in der jüngere Alters- bzw. Jahrganggruppen häufiger vertreten sind. Schließlich wurde die Befragungsmethodik aufgrund sinkender Antworttra-

ten erweitert. Zusätzlich zu den stets verwendeten schriftlichen Fragebögen wurde in der Erhebung des Jahres 2006 ein telefonischer Studienarm eingeführt; seit 2009 besteht für die Teilnehmer zusätzlich die Möglichkeit, den Fragebogen online auszufüllen.

Trotz dieser konzeptionellen Änderungen blieben der Kern des Surveys, die inhaltliche Schwerpunktsetzung und die Datenerhebung im größtmöglichen Ausmaß konstant, um langfristige Vergleiche und Trendanalysen durchführen zu können. Als wiederholte Querschnittsbefragung eignet sich die Studie somit für die Gewinnung wertvoller Informationen zu Prävalenz und Mustern des Substanzkonsums sowie zu Prävalenz und Trends des Umfangs von Missbrauch und Abhängigkeit (World Health Organization, 2000).

Im Rahmen der vorliegenden Arbeit werden Methodik und Design der aktuellen Erhebung des Epidemiologischen Suchtsurveys 2012 beschrieben. Es erfolgt ein Überblick über das Design und die Stichprobenziehung, die verwendeten Erhebungsinstrumente, die Durchführung der Befragung sowie die methodischen Analysen (Ausschöpfung, Gewichtung, Selektivitätsanalysen).

Methodik

Design und Stichprobenziehung

Zielpersonen des Epidemiologischen Suchtsurveys 2012 waren deutschsprachige und in Privathaushalten lebende Personen, die zwischen 1948 und 1994 geboren wurden. Die Grundgesamtheit beträgt somit ca. 51.63 Mio. Personen (Jahresdurchschnitt 2011, Statistisches Bundesamt). Die Ziehung der Personenstichprobe erfolgte auf Basis der Einwohnermelderegister in einem zweistufigen Zufallsverfahren. In einem ersten Schritt wurden basierend auf einer Kombination von Landkreisen und BIK-Gemeindegrößenklassen sowie unter Kontrolle der Regierungsbezirke und Bundesländer zufallsgesteuert Gemeinden bzw. Stadtteile von Großstädten ausgewählt. Die so ausgewählten 220 Gemeinden (251 Sample-Points) gewährleisteten die größtmögliche geographische und siedlungsstrukturelle Proportionalität der Stichprobe im Vergleich zur Grundgesamtheit. In einem zweiten Schritt wurden die Zielpersonen in einem systematischen Zufallsverfahren direkt aus den Einwohnermelderegistern der Primäreinheiten gezogen. Dabei erfolgte eine Stratifizierung nach Geschlecht und Geburtsjahrgang (1948–1952, 1953–1962, 1963–1972, 1973–1982, 1983–1987, 1988–1991, 1992–1994). Die Verteilung der Geburtsjahrganggruppen in der Stichprobe wurde disproportional zur Verteilung in der Grundgesamtheit gewählt, d. h. jüngere Geburtsjahrgänge wurden disproportional häufig und ältere Geburtsjahrgänge wurden disproportional selten gezogen. Es wurde eine Nettostichprobe von insgesamt 8000 Fällen angestrebt, sodass unter der Annahme einer Ausschöpfungsrate von 50 % und eines Anteils von 20 % neutralen Ausfällen etwa 80 Personenadressen pro Sample-Point gezogen

werden mussten. Aufgrund geringer Ausfälle in einigen Gemeinden betrug die Einsatzstichprobe letztlich insgesamt 20 006 Fälle.

Instrumente

Soziodemographie

Die Erfassung soziodemographischer Merkmale orientierte sich an den aktuellen demographischen Standards für persönlich-mündliche und schriftliche Befragungen des Statistischen Bundesamts (Statistisches Bundesamt, 2010). Neben grundlegenden Informationen zu Geschlecht und Geburtsjahr wurden folgende Aspekte erfragt: (a) Migrationshintergrund (Geburtsland und Staatsbürgerschaft des Probanden sowie der Eltern), (b) Familiensituation (Familienstand, Kinder, Haushaltsgröße), (c) Schulbildung, (d) Ausbildung, (e) Erwerbstätigkeit (Erwerbsstatus, berufliche Stellung), (f) Haushaltsnettoeinkommen.

Gesundheit

Der allgemeine Gesundheitszustand der Teilnehmer wurde mithilfe zweier 5-stufiger Ratings zur körperlichen und psychischen Gesundheit erfasst. Als Screening für psychische Störungen wurde die deutsche Version des K6 (Kessler et al., 2002) verwendet, welche die Häufigkeit unspezifischer und sozio-emotionalen Distresses in den letzten 30 Tagen erfasst. Darüber hinaus wurden die Probanden gebeten anzugeben, ob ein Arzt bei ihnen jemals eine oder mehrere chronische Krankheiten wie z. B. Allergien, Haut-, Atemwegs- oder Herz-Kreislauf-Erkrankungen festgestellt hat.

Substanzkonsum

Alkohol

Informationen zur durchschnittlichen Menge des Alkoholkonsums wurden über einen getränkespezifischen Menge-Frequenz-Index für die letzten 30 Tage bzw. die letzten 12 Monate erhoben (Gmel & Rehm, 2004). Hierfür wurden für Bier, Wein/Sekt, Spirituosen sowie alkoholhaltige Mixgetränke die Häufigkeit des Konsums sowie die durchschnittliche Konsummenge pro Trinktag erfragt. Basierend auf getränkespezifischen Alkoholgehalten (Bühlinger et al., 2000) wurde die Gesamtmenge reinen Alkohols in Gramm berechnet. Zusätzlich zur Konsummenge wurde das Einstiegsalter in den (regelmäßigen) Konsum sowie die Häufigkeit des Betrunkenseins und des episodischen Rauschtrinkens (5 oder mehr Gläser Alkohol an einem Tag) erfasst.

Tabak

Neben der Prävalenz des Rauchens von Zigaretten/Zigarren/Zigarillos/Pfeifen (Lebenszeit, 12 Monate, 30 Tage) wurde der Konsum von Shisha (Wasserpfeife) und das starke Rauchen von täglich im Durchschnitt 20 oder mehr Zigaretten (oder 10 oder mehr Zigarillos oder 7 oder mehr Pfeifen oder 5 oder mehr Zigarren) in den letzten 12 Monaten erfragt. Die durchschnittliche Konsummenge wurde für Zigaretten sowie Zigarren/Zigarillos/Pfeifen über einen Menge-Frequenz-Index erfasst, in den Informationen zur Anzahl der Konsumtage in den letzten 30 Tagen sowie zur durchschnittlichen Konsummenge pro Konsumtag eingegangen. Das Einstiegsalter wurde für das Rauchen und das tägliche Rauchen erfragt. Darüber hinaus wurden die Teilnehmer nach Aufhörversuchen und deren Erfolgen gefragt.

Drogen

Das Konsumverhalten von Drogen wurde getrennt für Cannabis (Haschisch, Marihuana), Aufputzmittel/Amphetamine, Ecstasy, LSD, Heroin, andere Opiate, Kokain, Crack, Schnüffelstoffe, Pilze sowie Spice/Smoke/Space/Badesalze/Cathinone o.a. erfragt. Für alle Substanzen wurden die Prävalenz und Frequenz bezogen auf die Lebenszeit, die letzten 12 Monate und 30 Tage sowie das Erstkonsumalter erfasst.

Medikamente

Informationen zum Gebrauch von Medikamenten wurden für Schmerzmittel (Analgetika), Schlafmittel (Hypnotika), Beruhigungsmittel (Tranquilizer), Anregungsmittel (Analeptika), Appetitzügler (Anorektika), Antidepressiva, Neuroleptika (Antipsychotika) und Anabolika (Anabole Steroide) erfasst. Um den Befragten die Zuordnung spezifischer Präparate zu diesen Gruppen zu erleichtern, wurde eine Liste der gebräuchlichsten Medikamente vorgegeben. Für alle Substanzgruppen wurde die 12-Monats- und 30-Tage-Prävalenz sowie die 30-Tage-Frequenz erhoben. Darüber hinaus wurde erfasst, ob die im letzten Monat eingenommenen Medikamente vom Arzt verordnet waren oder nicht.

Substanzbezogene Störungen

Substanzbezogene Störungen wurden mithilfe des Münchener Composite International Diagnostic Interview (M-CIDI; Wittchen et al., 1995) für Alkohol, Tabak, Cannabis, Kokain, Amphetamine, Schmerzmittel, Schlafmittel und Beruhigungsmittel erfasst. Für alle Substanzen wurden die Kriterien für die Diagnosen nach DSM-IV und DSM-5 erhoben.

Inanspruchnahme von Hilfe

Zur Erfassung der Inanspruchnahme von Hilfe wurden die Teilnehmer zunächst gefragt, ob der Konsum von Alkohol, illegalen Drogen oder Medikamenten eine schädliche Wirkung auf ihre Gesundheit oder ihr Leben hatte. Darüber hinaus wurde eine Liste von möglichen Hilfsangeboten vorgegeben, deren Inanspruchnahme innerhalb der letzten 12 Monate angegeben werden sollte (Hausarzt, Psychotherapie, ambulante Suchtberatung, stationäre Entgiftung, stationäre Rehabilitation, Selbsthilfegruppe, Unterstützung durch Freunde/Partner/Familie). Für den Bereich des Tabakkonsums wurden die Befragten gebeten anzugeben, welche konkreten Hilfsmittel sie bei ihrem letzten Aufhörversuch innerhalb der letzten zwei Jahre verwendet hatten (Entwöhnungskurs oder -seminar, nikotinhaltige Präparate (Pflaster, Kaugummi, Lutschtabletten, Sublingualtabletten, Nikotin-Inhaler), verschreibungspflichtige Medikamente (Bupropion, Vareniclin), Bücher/Broschüren/CDs/Videos/DVDs, Hypnose, Akupunktur/Akupressur, computergestützte Interventionen, elektronische Zigarette, ärztliche Beratung, Rauchtelefon, Unterstützung durch Partner/Freunde/Familie, andere Hilfsmittel, ohne Hilfsmittel oder Hilfen).

Schäden durch den Alkoholkonsum Dritter

Die Teilnehmer wurden gefragt, ob es in den letzten 12 Monaten zu Gefährdungen oder Schädigungen durch den Alkoholkonsum dritter Personen gekommen ist. Erfasst wurden familiäre oder Eheprobleme, Autofahrten mit einer alkoholisierten Person, Autounfälle, finanzielle Probleme, körperliche Verletzungen, Zerstörung von Kleidung oder Gegenständen, Beschimpfungen oder Beleidigungen, nächtliche Störungen, Belästigungen auf öffentlichen Plätzen, Belästigungen im privaten Rahmen sowie Verängstigung durch betrunkenen Personen. Bei den eingesetzten Fragen handelt es sich um die deutsche Übersetzung eines in Dänemark eingesetzten Instruments (Bloomfield & Grittner, 2012).

Non-response-Fragebogen

Alle Personen, die nicht bereit waren, an der Studie teilzunehmen, oder von denen keine Rückmeldung vorlag, wurden gebeten, acht Non-response-Fragen zu beantworten. Erfasst wurden dabei (a) die Lebenszeit-, 12-Monats- und 30-Tage-Prävalenz des Alkoholkonsums, (b) die Häufigkeit des episodischen Rauschtrinkens in den letzten 30 Tagen, (c) der Menge-Frequenz-Index des Zigarettenkonsums, (d) die Lebenszeit- und 12-Monats-Prävalenz des Cannabiskonsums sowie (e) die 12-Monats-Prävalenz des Gebrauchs von Schmerz- und Schlafmitteln.

Durchführung

Die Befragungen des Epidemiologischen Suchtsurveys 2012 wurden im Methodenmix durchgeführt, d.h. die Daten wurden schriftlich, telefonisch und im Internet erhoben. Die Feldarbeiten wurden in der Zeit von April bis August 2012 vom infas Institut für angewandte Sozialwissenschaft GmbH durchgeführt. In einem ersten Schritt wurde die Gesamtstichprobe in zwei Arme geteilt, je nachdem, ob für die gezogene Adresse eine Telefonnummer ermittelt werden konnte (telefonischer Arm) oder nicht (schriftlicher Arm). Die Telefonnummern-Recherche erfolgte dabei in aktuellen und ggf. älteren digitalen Telefonnummernverzeichnissen.

Zielpersonen des telefonischen Arms erhielten zunächst eine schriftliche Einladung für die Studie zusammen mit einem Informationsschreiben, einer Datenschutzerklärung und einem Begleitschreiben des Bundesministeriums für Gesundheit (BMG). Die telefonische Kontaktaufnahme erfolgte durch geschulte Interviewer. Bei den Gesprächen wurde den Zielpersonen, die sich nicht telefonisch befragen lassen wollten, die Beantwortung des Fragebogens in schriftlicher Form bzw. im Internet angeboten. Jede Person, die erreicht wurde aber nicht an der Studie teilnehmen wollte, wurde um die Beantwortung des Non-response-Fragebogens gebeten. Zielpersonen, deren ermittelte Telefonnummer ungültig war oder die auch nach 20 Kontaktversuchen nicht erreicht werden konnten, wurde der schriftliche Fragebogen und ein Zugangs-Code für die Internet-Befragung postalisch zugesandt. Verbleibende Personen, die bis zum Feldende nicht teilgenommen hatten, erhielten den schriftlichen Non-response-Fragebogen mit der Bitte um Rücksendung.

Im schriftlichen Arm erhielten die Zielpersonen zusammen mit dem Anschreiben und den beigefügten Unterlagen den schriftlichen Fragebogen zum Selbstausfüllen. Im Anschreiben wurde auch die Teilnahme im Internet angeboten. Dazu wurden die Internetadresse für die Befragung und der persönliche Zugangs-Code angegeben. Alle Personen, die sich nicht an der Studie beteiligten, erhielten maximal drei Erinnerungsschreiben, in denen auch der Online-Zugangs-Code erneut übermittelt sowie eine Beantwortung am Telefon angeboten wurde. Erfolgte bis zum Feldende keine Rückmeldung, wurde der Non-response-Fragebogen versendet.

Analysen

Datenprüfung und Auswertung

Nach Eingang aller Daten wurde zunächst eine manuelle Prüfung der Vollständigkeit und Plausibilität durchgeführt. Je nach Befragungsmodus erfolgte die Datenerfassung entweder maskengestützt durch professionelle KodiererInnen am Computer (schriftliche Fragebögen) bzw. direkt in die integrierte Software (telefonische Interviews und Online-Fragebögen). Schließlich erfolgte eine manuelle

Prüfung auf Filterfehler, Plausibilität und Inkonsistenzen und ggf. eine Bereinigung entsprechender Auffälligkeiten.

Bei allen Analysen wurde das komplexe Stichproben-Design berücksichtigt. Dies beinhaltet sowohl das mehrstufige Auswahlverfahren (Clusterung) als auch die Gewichtung der Daten. Für die Auswertungen wurde das Programm Stata 12.1 SE (Stata Corp LP, College Station, TX) genutzt. Zur Varianzschätzung kamen Taylorreihen zum Einsatz (Groves et al., 2009).

Ausschöpfung

Für die Berechnung der Ausschöpfungsrates wurde zunächst der Anteil der nicht befragbaren Personen ermittelt. Diese nicht realisierten Befragungen lassen sich in drei Gruppen aufteilen. Systematische Ausfälle sind prinzipiell für die Studie verfügbare Personen, die sich aber nicht beteiligt haben. Dazu zählen Zielpersonen, die (a) den Fragebogen zurücksenden oder online ausfüllen wollten bzw. den Telefontermin nicht eingehalten haben, (b) in der Feldzeit nicht erreichbar waren, (c) aus gesundheitlichen Gründen nicht befragbar waren oder (d) die Teilnahme schriftlich oder telefonisch verweigert hatten. Im Gegensatz dazu war es für neutrale Ausfälle niemals möglich, in die Stichprobe der Studie aufgenommen zu werden. Dies betrifft Zielpersonen, die (a) unbekannt verzogen waren, (b) verstorben waren, (c) nicht deutschsprachig waren, (d) nicht in der Altersgruppe des Surveys waren oder für die (e) die ermittelte Telefonnummer ungültig war. Eine dritte Gruppe nicht realisierter Befragungen stellen Zielpersonen dar, von denen zum Ende der Feldzeit überhaupt keine Information vorlag (Zielpersonen ohne Rücklauf). Der Prozentsatz der realisierten Befragungen an der Einsatzstichprobe unter Ausschluss aller nicht realisierten Interviews ergibt die Bruttorealisierung. Für die Berechnung der Nettoausschöpfung wurde die Einsatzstichprobe um den Anteil neutraler Ausfälle bereinigt. Da dieser Anteil für die Gruppe der Zielpersonen ohne Rücklauf nicht bekannt ist, erfolgte eine Schätzung auf Basis der Personen mit bekanntem Rücklaufstatus, d.h. der Prozentsatz der neutralen Ausfälle an allen Zielpersonen mit Rücklauf (systematische + neutrale Ausfälle) wurde ebenfalls für die Gruppe der Zielpersonen ohne Rücklauf verwendet (Biemer & Lyberg, 2003). Die Nettoausschöpfung ergibt sich letztlich als das Verhältnis der realisierten Stichprobe an der bereinigten Einsatzstichprobe.

Gewichtung

Eine Gewichtung der Daten ist notwendig, um sicherzustellen, dass die realisierte Stichprobe hinsichtlich zentraler Merkmale mit der Grundgesamtheit vergleichbar ist, und um somit bevölkerungsrepräsentative Aussagen treffen zu können. Zunächst wurde die disproportionale Verteilung der Stichprobe nach Geburtsjahrgängen, die aus dem Studien-Design resultierte, durch Bildung eines Design-Ge-

wichts an die Verteilung der Grundgesamtheit angepasst. Dieser Gewichtungsfaktor wurde proportional zur inversen Auswahlwahrscheinlichkeit auf der jeweiligen Auswahlstufe gebildet. Zusätzlich wurde ein Redressement-Gewicht berechnet, welches die Stichprobe bezüglich verschiedener externer Variablen an die Grundgesamtheit anpasst. Als Anpassungs-Merkmale wurden dabei Geschlecht, Geburtsjahrgang, höchster Schulabschluss sowie Bundesland und BIK-Gemeindegrößenklasse herangezogen. Die Randverteilungen der Grundgesamtheit wurden dem Mikrozensus 2011 entnommen. Für die Gewichtung wurde ein Iterative Proportional Fitting Algorithmus verwendet (Gelman & Carlin, 2001).

Zur Bewertung der Gewichtung wurden zwei Kennwerte berechnet, die effektive Fallzahl n' und das sog. Effektivitätsmaß E (Gabler, Hoffmeyer-Zlotnik & Krebs, 1994; Little, Lewitzky, Heeringa, Lepkowski & Kessler, 1997). Gewichtung resultiert in der Verringerung der Stichprobengröße von n auf n' . Die effektive Fallzahl wird über die Formel $n' = (\sum g_i)^2 / \sum g_i^2$ berechnet, wobei g die Gewichtungsvariable ist. Das Effektivitätsmaß basiert auf der Varianz des Gewichtungsfaktors, d.h. eine größere Varianz hat einen größeren Einfluss auf den Stichprobenfehler. Die Effektivität kann als der Prozentwert $100 * n'/n$ berechnet werden. Je näher E am Wert 100 % liegt, desto geringer ist die gewichtungsbedingte Reduktion der Fallzahl. Neben der effektiven Fallzahl und dem Effektivitätsmaß wird jeweils das Minimum und das Maximum der Gewichtungsfaktoren sowie deren Standardabweichung berichtet.

Selektivitätsanalysen

Zur Beurteilung der Generalisierbarkeit der Ergebnisse ist es notwendig, Verzerrungen aufgrund von Ausfällen und unterschiedlicher Erhebungsmethodik zu kennen und ihre Auswirkungen auf relevante Merkmale abzuschätzen. Im Rahmen des Epidemiologischen Suchtsurveys 2012 wurden drei Arten von Selektivitätsanalysen durchgeführt.

Vergleich von Einsatz- und realisierter Stichprobe mit der Grundgesamtheit

Ein deskriptiver Vergleich der bereinigten Bruttostichprobe mit der Grundgesamtheit war hinsichtlich der Schichtungsvariablen (z.B. Bundesland, BIK-Gemeindegrößenklasse) sowie der grundlegenden soziodemographischen Variablen Alter und Geschlecht möglich. Darüber hinaus wurde die gewichtete realisierte Stichprobe hinsichtlich weiterer zentraler soziodemographischer Variablen mit der erwachsenen Allgemeinbevölkerung verglichen. Dafür wurden die Merkmale Staatsbürgerschaft, Familienstand, Schulabschluss und Haushaltsnettoeinkommen verwendet. Daten zur Verteilung dieser Variablen in der Grundgesamtheit wurden vom Statistischen Bundesamt bereitgestellt.

Non-response-Analysen

Auf Basis des Non-response-Fragebogens wurde untersucht, inwieweit systematische Verzerrungen aufgrund eines Zusammenhangs zwischen den Untersuchungsmerkmalen und der Teilnahmebereitschaft vorliegen. Neben einem Vergleich soziodemographischer Variablen wurden die im Rahmen der Non-response-Studie erhobenen Konsumvariablen analysiert. Dafür wurden logistische bzw. lineare Regressionsanalysen mit den Konsumvariablen als Kriterium und der Teilnahmebereitschaft als Prädiktor berechnet. Diese Analysen wurden nach Alter, Geschlecht, Bundesland und Erhebungsmethodik adjustiert.

Modus-Analysen

Im Rahmen von Modus-Analysen wurde geprüft, ob die unterschiedlichen Erhebungsmethoden (schriftlich, Telefon, Internet) einen Einfluss auf den selbstberichteten Substanzkonsum hatten. Untersucht wurde der Konsum von Alkohol (30-Tage-Prävalenz des Konsums und episodischen Rauschtrinkens), Tabak (30-Tage-Prävalenz, durchschnittliche Anzahl Zigaretten pro Tag in den letzten 30 Tagen), Cannabis (Lebenszeit- und 12-Monats-Prävalenz) sowie der Gebrauch von Medikamenten (12-Monats-Prävalenz in Bezug auf Schmerzmittel und Schlafmittel). Es wurden logistische bzw. lineare Regressionen adjustiert nach Alter, Geschlecht, Bundesland, Schulbildung und Haushaltsnettoeinkommen berechnet. Die schriftliche Befragung wurde als Referenzkategorie genutzt.

Ergebnisse

Realisierte Stichprobe

Im Rahmen des Epidemiologischen Suchtsurveys 2012 konnten insgesamt 9099 Interviews realisiert werden. Davon nahmen 4516 Befragte (49.6 %) schriftlich, 3898 Befragte (42.8 %) telefonisch und 685 Befragte (7.5 %) im Internet teil.

Die Stichproben-Realisierung in den beiden Studienarmen ist in Abbildung 1 dargestellt. Für 47.5 % ($n = 9499$) der insgesamt 20 006 eingesetzten Adressen konnte eine Telefonnummer recherchiert werden (telefonischer Arm), die verbleibenden 52.5 % ($n = 10 507$) erhielten den Fragebogen postalisch (schriftlicher Arm). In der Einsatzstichprobe ohne Telefonnummer haben insgesamt 60.8 % der kontaktierten Personen nicht an der Studie teilgenommen. Der weitaus größte Teil (54.1 %) entfiel dabei auf die Kategorie „ohne Rücklauf“, d.h. von diesen Personen lag auch nach drei Erinnerungsschreiben keine Rückmeldung vor. Von den 4119 Teilnehmern haben 86.2 % den schriftlichen Fragebogen ausgefüllt, 13.5 % nutzten das Angebot der Online-Beantwortung und 0.3 % beantworteten den Fragebogen am Telefon. In der Einsatzstichprobe

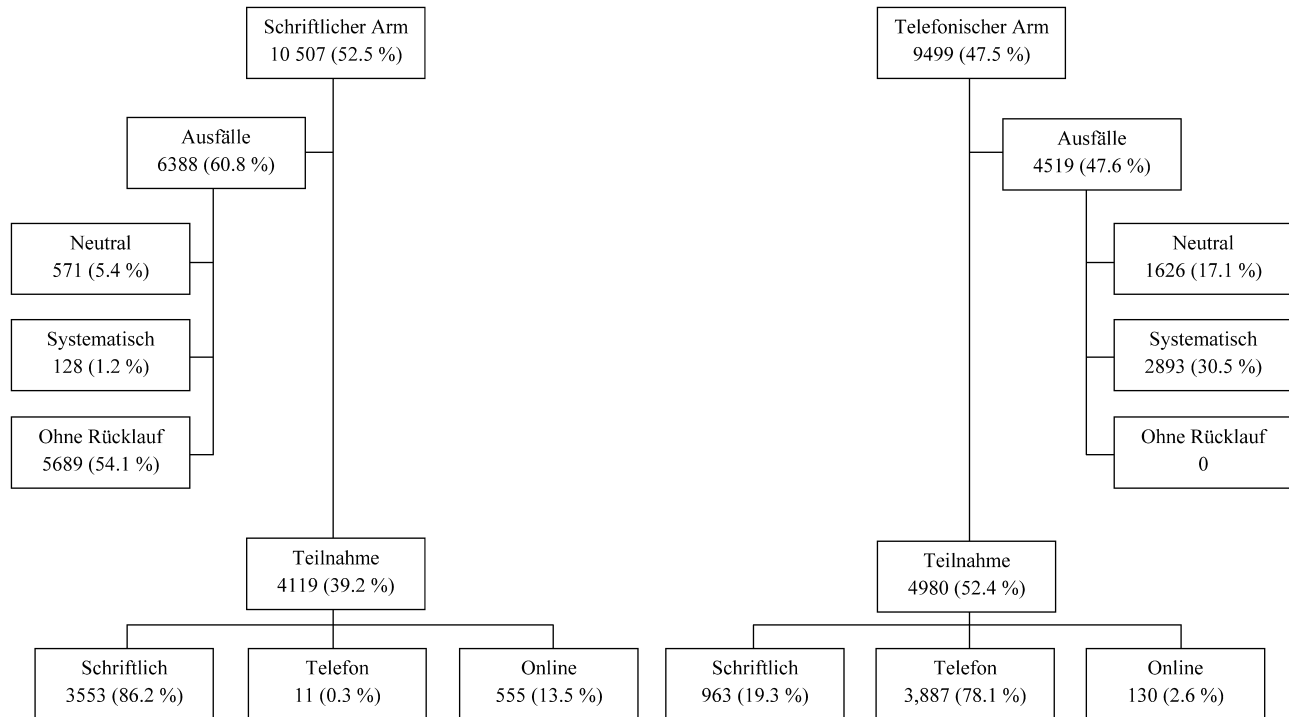


Abbildung 1. Stichproben-Realisierung des ESA 2012 nach Studienarm.

mit Telefonnummer konnten 4980 Interviews realisiert werden. Davon wurden 78.1 % telefonische Befragungen durchgeführt, die verbleibenden Fälle wurden schriftlich (19.3 %) oder online (2.6 %) realisiert. Bei den 47.6 % Ausfällen in diesem Arm handelte es sich zum überwiegenden Teil (30.5 %) um systematische Ausfälle.

Die Einsatzstichprobe für die Non-response-Befragung umfasste insgesamt 10 123 Personen, 5695 (56.3 %) aus dem schriftlichen und 4428 (43.7 %) aus dem telefonischen Arm. Von der Non-response-Befragung ausgeschlossen waren unter anderem Personen, die verstorben, unbekannt verzogen oder aus gesundheitlichen Gründen nicht befragbar waren, und Personen, die den telefonischen Kontakt bereits beendet hatten, bevor der Interviewer um die Beantwortung der Non-response Fragen bitten konnte. Insgesamt beantworteten 758 Personen (7.5 %) die acht Non-response-Fragen. Die Beteiligung lag in der Einsatzstichprobe ohne Telefonnummer (9.3 %) deutlich höher als in der Einsatzstichprobe mit Telefonnummer (5.1 %).

Ausschöpfung

Insgesamt 10 907 Personen (54.5 %) der Einsatzstichprobe konnten nicht befragt werden (Tabelle 1). Von diesen waren 3021 (27.7 %) systematische Ausfälle, 2197 neutrale Ausfälle (20.1 %) und 5689 Zielpersonen ohne Rücklauf (52.2 %). Eine Aufgliederung der systematischen Ausfälle ergab 49.1 % Verweigerungen, 43.0 % nicht erreichbare Zielpersonen, 5.4 % Zielpersonen, die angaben, den Fragebogen zurücksenden zu wollen, bzw. Termine nicht ein-

hielten, und 2.5 % Zielpersonen, die wegen Krankheit nicht befragt werden konnten. Die neutralen Ausfälle unterteilten sich in 59.9 % ungültige Telefonnummern, 36.3 % unbekannt verzogene Empfänger, 2.5 % nicht deutschsprachige Zielpersonen, 1.0 % verstorbene Zielpersonen und 0.4 % Zielpersonen außerhalb des Altersbereichs der Studie. Die Summe aus realisierten Befragungen, systematischen und neutralen Ausfällen ergab insgesamt 14 317 Personen mit bekanntem Rücklaufstatus. Der Anteil stichprobenneutraler Ausfälle an diesen Personen lag bei 15.3 % (n = 2197). Der gleiche Prozentsatz neutraler Ausfälle wurde für die Zielpersonen ohne Rücklauf angenommen und ergab 873 geschätzte zusätzliche neutrale Ausfälle. Die um alle neutralen Ausfälle bereinigte Einsatzstichprobe umfasste somit 16 936 Personen (84.7 %). Von den realisierten 9099 Interviews mussten 15 Fälle (0.1 %) ausgeschlossen werden, da die Fragebögen nicht auswertbar waren. Die endgültige Stichprobengröße des Surveys liegt somit bei n = 9084 Personen.

Auf Basis dieser Informationen ergibt sich eine Bruttorealisierung von 45.4 %. Die Nettoausschöpfung nach Ausschluss neutraler Ausfälle beträgt 53.6 %. Es zeigen sich zum Teil deutliche Unterschiede in der Ausschöpfung nach Region sowie Alter und Geschlecht (Tabelle 2). Mit 41.0 % ist die Antwortrate in Mecklenburg-Vorpommern am niedrigsten und mit 56.4 % in Bayern am höchsten. Frauen nahmen insgesamt zu einem größeren Anteil an der Studie teil als Männer (55.4 % vs. 46.7 %). Die Ausschöpfungsraten sind bei den älteren Geburtsjahrgängen (1948 bis 1962) höher als in mittleren und jüngeren Geburtsjahrgängen. Am niedrigsten ist die Teilnehmerate bei

Tabelle 1
Rücklaufstatus und Ausschöpfung

	n	Prozent
Einsatzadressen insgesamt	20 006	100.0
Nicht befragbare Personen	10 907	54.5
(1) Systematische Ausfälle	3021	27.7
(2) Neutrale Ausfälle	2197	20.1
(3) Zielpersonen ohne Rücklauf	5689	52.2
Zielpersonen mit bekanntem Rücklauf ^{a)}	14 317	71.6
Neutrale Ausfälle	2197	15.3
Geschätzte neutrale Ausfälle an Zielpersonen ohne Rücklauf ^{b)}	873	15.3
Bereinigte Einsatzstichprobe	16 936	84.7
Realisierte Interviews	9099	
Ausschluss nach Datenprüfung	15	
Bruttorealisierung	9084	45.4
Nettoausschöpfung	9084	53.6

Anmerkungen. ^{a)} Teilnehmer (n = 9099), systematische Ausfälle (n = 3021) und neutrale Ausfälle (n = 2197); ^{b)} Schätzung auf Basis der Personen mit bekanntem Rücklaufstatus, d. h. der Prozentsatz der neutralen Ausfälle an allen Zielpersonen mit Rücklauf wird ebenfalls für die Gruppe der Zielpersonen ohne Rücklauf verwendet.

Männern, die zwischen 1983 und 1987 geboren wurden (35.5 %), am höchsten ist sie bei zwischen 1953 und 1962 geborenen Frauen (61.9 %).

Gewichtung

Das Design-Gewicht weist ein Minimum von 0.38 und ein Maximum von 1.40 auf. Die Standardabweichung beträgt 0.32. Das Effektivitätsmaß E beträgt 90.8 %, was einer effektiven Fallzahl von $n' = 8249$ Fällen entspricht. Da das Design-Gewicht im Redressement-Gewicht enthalten ist und die beschriebenen Anpassungsmerkmale zusätzlich in die Gewichtung einbezogen werden, wirkt sich das Redressement-Gewicht zwangsläufig stärker auf die effektive Fallzahl aus. Diese liegt für die verwendete Gewichtungsvariable bei $n' = 6142$, die Effektivität beträgt 67.6 %. Der Range des Gewichtungsfaktors liegt zwischen 0.16 und 6.87, die Standardabweichung beträgt 0.69.

Vergleiche mit der Grundgesamtheit

Die bereinigte Einsatzstichprobe zeigt eine sehr gute regionale Vergleichbarkeit zur gesamtdeutschen Bevölkerung (Tabelle 2). Männer sind insgesamt leicht unter- und Frauen entsprechend leicht überrepräsentiert. Aufgrund des disproportionalen Studiendesigns sind in der Einsatzstichprobe bei beiden Geschlechtern ältere Geburtsjahrgänge zu einem niedrigeren und jüngere Geburtsjahrgänge zu einem höheren Prozentsatz vertreten. Einen Vergleich der ungewichteten und gewichteten realisierten Stichprobe mit der Grundgesamtheit zeigt Tabelle 3. In der ungewichteten Stichprobe ist der Anteil lediger Personen und

solcher mit Abitur höher als in der erwachsenen Allgemeinbevölkerung. Durch die Gewichtung konnte diese Verteilung jedoch an die Grundgesamtheit angeglichen werden. Gleichzeitig sinkt in der gewichteten Stichprobe jedoch die Übereinstimmung mit der Grundgesamtheit für Personen mit Schulabschluss „mittlere Reife/Polytechnische Oberschule“ und verheiratete Personen. Grundsätzlich in der Stichprobe der Studie unterrepräsentiert sind nicht-deutsche Staatsangehörige. Bezüglich des Haushaltsnettoeinkommens zeigen sich keine bedeutsamen Unterschiede.

Non-response-Effekte

Bei allen erfragten Substanzen zeigen sich Unterschiede im selbstberichteten Konsumverhalten zwischen Teilnehmern und Nicht-Teilnehmern, die den Non-response-Fragebogen beantworteten (Tabelle 4). So berichteten Nicht-Teilnehmer seltener, in den letzten 30 Tagen Alkohol sowie in ihrem Leben und in den letzten 12 Monaten Cannabis konsumiert zu haben. Im Gegensatz dazu zeigen Nicht-Teilnehmer eine höhere 12-Monats-Prävalenz des Schlafmittelgebrauchs. Für Tabak und Schmerzmittel konnten keine Unterschiede in der Prävalenz festgestellt werden. Unter den Personen, die sich nicht am Survey beteiligen wollten, war die durchschnittliche Anzahl der gerauchten Zigaretten höher als bei Teilnehmern. Außerdem hatte ein größerer Prozentsatz im letzten Monat mindestens einmal fünf oder mehr Gläser Alkohol zu einer Gelegenheit getrunken.

Tabelle 2

Regionale und demographische Verteilung der Einsatzstichprobe, der realisierten Stichprobe und der Grundgesamtheit sowie Bruttorealisierung des ESA 2012 (Prozent)

	Einsatzstichprobe (bereinigt)	Bruttorealisierung	Realisierte Stichprobe (gewichtet)	Grundgesamtheit ^{a)}
Bundesland				
Schleswig-Holstein	3.0	49.3	3.4	3.4
Hamburg	2.5	47.4	2.3	2.3
Niedersachsen	9.6	55.4	9.5	9.4
Bremen	0.8	50.0	0.8	0.8
Nordrhein-Westfalen	22.1	49.1	21.7	21.6
Hessen	7.2	50.3	7.5	7.4
Rheinland-Pfalz	4.7	47.7	4.9	4.8
Baden-Württemberg	13.3	53.1	13.2	13.1
Bayern	15.2	56.4	15.5	15.3
Saarland	1.2	50.5	1.4	1.3
Berlin	4.3	42.0	4.5	4.5
Brandenburg	3.2	48.5	3.1	3.2
Mecklenburg-Vorpommern	1.5	41.0	2.3	2.1
Sachsen	5.3	52.7	4.9	5.1
Sachsen-Anhalt	3.2	49.8	2.8	2.9
Thüringen	3.0	46.3	2.7	2.8
Jahrgangsgruppe * Geschlecht				
Männer				
Gesamt	47.5	46.7	50.8	50.6
1948–1952	4.7	59.3	4.9	5.7
1953–1962	8.4	56.0	12.0	11.9
1963–1972	8.7	46.8	13.1	13.0
1973–1982	8.0	36.9	9.5	9.5
1983–1987	5.9	35.5	5.0	4.9
1988–1991	5.8	42.9	3.9	3.9
1992–1994	6.1	51.0	2.5	1.7
Frauen				
Gesamt	52.5	55.4	49.2	49.4
1948–1952	5.6	60.4	5.1	5.9
1953–1962	9.7	61.9	11.9	11.8
1963–1972	9.2	57.3	12.5	12.4
1973–1982	9.4	52.5	9.0	9.3
1983–1987	6.2	48.4	4.7	4.7
1988–1991	6.2	50.5	3.8	3.7
1992–1994	6.3	54.0	2.3	1.6

Anmerkungen. ^{a)} Jahresdurchschnittswerte 2011.

Modus-Effekte

Auch hinsichtlich der Befragungsmethoden zeigen sich Unterschiede im Konsumverhalten (Tabelle 5). Verglichen mit schriftlich befragten Probanden wiesen telefonisch befragte Teilnehmer eine niedrigere 30-Tage-Prävalenz des Tabakkonsums, eine niedrigere Lebenszeit- und 12-Monats-Prävalenz des Cannabiskonsums sowie eine niedrigere 12-Monats-Prävalenz des Schlafmittelgebrauchs auf. Personen, die den Fragebogen online ausfüllten, zeigen ebenfalls niedrigere Prävalenzwerte als schriftlich Befrag-

te, und zwar im Hinblick auf die 30-Tage-Prävalenz des Alkoholkonsums sowie die Lebenszeit- und 12-Monats-Erfahrung mit Cannabis.

Diskussion

Wie in den vorangegangenen Erhebungen steht im Epidemiologischen Suchtsurvey 2012 die Untersuchung von Konsumverhalten und Störungen im Zusammenhang mit dem Gebrauch von legalen und illegalen Substanzen sowie

Tabelle 3

Verteilung soziodemographischer Merkmale in der ungewichteten und gewichteten Stichprobe und in der Grundgesamtheit (Prozent)

	Ungewichtete Stichprobe			Gewichtete Stichprobe			Bevölkerung ^{a)} (18- bis 64-Jährige)		
	Gesamt	Männer	Frauen	Gesamt	Männer	Frauen	Gesamt	Männer	Frauen
Staatsbürgerschaft									
deutsch	94.1	94.4	93.8	93.8	93.5	94.1	88.9	88.9	89.0
andere	5.9	5.6	6.2	6.2	6.5	5.9	11.1	11.1	11.0
keine Angabe	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Familienstand									
ledig	42.2	47.0	38.5	34.3	39.0	29.4	39.4	43.9	34.9
verheiratet	48.5	45.3	51.0	54.6	51.8	57.6	48.2	45.6	51.0
verwitwet	1.6	0.9	2.1	1.7	1.0	2.5	2.0	0.8	3.1
geschieden	6.4	5.6	7.0	7.7	6.6	8.8	10.4	9.7	11.1
keine Angabe	1.3	1.2	1.4	1.7	1.7	1.7	0.0	0.0	0.0
Schulabschluss									
in Schulausbildung	3.2	3.3	3.2	1.7	1.6	1.7	1.8	1.8	1.8
Hauptschulabschluss	15.5	18.2	13.4	32.8	37.1	28.4	28.0	30.4	25.6
Mittlere Reife/Polytechnische Oberschule	32.2	30.1	33.8	29.9	26.7	33.1	34.0	31.0	37.1
(Fach-) Abitur	45.9	44.8	46.7	31.0	29.3	32.7	32.0	32.7	31.3
ohne Abschluss	1.3	1.8	1.0	2.9	3.7	2.0	3.7	3.6	3.8
anderer Abschluss	1.4	1.5	1.4	1.4	1.4	1.4	0.2	0.2	0.2
keine Angabe	0.5	0.4	0.5	0.4	0.3	0.5	0.2	0.2	0.2
Haushaltsnettoeinkommen ^{b)}									
unter 500 €	2.2	2.3	2.1	2.1	2.3	1.9	1.5	1.8	1.2
500–2000 €	29.1	26.6	31.0	32.2	30.6	33.7	29.3	28.1	30.5
2000–4000 €	41.6	42.4	41.0	43.3	44.6	41.9	38.6	38.7	38.4
4000 € und mehr	18.4	20.6	16.8	15.2	16.1	14.3	17.2	17.8	16.7
keine Angabe	8.7	8.2	9.0	7.2	6.5	8.0	13.4	13.6	13.3

Anmerkungen. ^{a)} Jahresdurchschnittswerte 2011; ^{b)} aufgrund abweichender Einkommensgrenzen beim Statistischen Bundesamt ist nur eine grobe Kategorisierung des Haushaltsnettoeinkommens möglich.

Tabelle 4

Adjustierte Konsumvariablen nach Teilnahmebereitschaft, n (%)^{a)}

	Teilnehmer	Nicht-Teilnehmer ^{b)}
	n = 9084	n = 758
Alkoholkonsum		
30-Tage-Prävalenz	6751 (74.7)	491 (65.5)**
Episodisches Rauschtrinken ^{c)}	2357 (35.2)	260 (54.2)**
Tabakkonsum		
30-Tage-Prävalenz	2315 (25.7)	184 (24.6)
Durchschnittliche Anzahl Zigaretten pro Tag, M (SD) ^{c)}	11.13 (9.90)	13.01 (10.89)*
Cannabiskonsum		
Lebenszeitprävalenz	2159 (23.9)	121 (16.1)**
12-Monats-Prävalenz	505 (5.6)	23 (3.1)*
Medikamentengebrauch letzte 12 Monate		
Schmerzmittel	5590 (62.0)	451 (59.7)
Schlafmittel	471 (5.2)	54 (7.2)**

Anmerkungen. ^{a)} Die Adjustierung nach Alter, Geschlecht, Bundesland und Interviewart wurde mit logistischen bzw. linearen Regressionsmodellen vorgenommen; ^{b)} Personen, die den Non-response Fragebogen beantwortet haben (7.5 % aller Nicht-Antworter); ^{c)} Bezogen auf 30-Tage-Konsumenten; * p < .05, **p < .01 für Vergleich mit „Teilnehmer“.

Tabelle 5
Adjustierte Konsumvariablen nach Art der Methode, n (%)^{a)}

	Schriftlich n = 4505	Telefon n = 3895	Internet n = 684
Alkoholkonsum			
30-Tage-Prävalenz	3316 (72.8)	2951 (73.5)	484 (65.1)**
Episodisches Rauschtrinken ^{b)}	1134 (35.9)	1031 (34.5)	192 (38.8)
Tabakkonsum			
30-Tage-Prävalenz	1265 (33.5)	880 (25.8)**	170 (31.1)
Durchschnittliche Anzahl Zigaretten pro Tag, M (SD) ^{b)}	13.18 (0.42)	12.45 (0.39)	12.12 (0.92)
Cannabiskonsum			
Lebenszeitprävalenz	1306 (28.7)	648 (15.0)**	205 (28.2)*
12-Monats-Prävalenz	293 (6.0)	168 (2.3)**	44 (4.1)**
Medikamentengebrauch letzte 12 Monate			
Schmerzmittel	2845 (63.9)	2334 (59.9)	411 (59.6)
Schlafmittel	275 (6.6)	170 (4.5)*	26 (3.8)

Anmerkungen. ^{a)} Die Adjustierung nach Alter, Geschlecht, Bundesland, Schulbildung und Haushaltsnettoeinkommen wurde mit logistischen bzw. linearen Regressionsmodellen vorgenommen; ^{b)} Bezogen auf 30-Tage-Konsumenten; * p < .05, **p < .01 für Vergleich mit „Schriftlich“.

Medikamenten im Mittelpunkt. Der Fragebogen ist modular aufgebaut, um den Zielen eines konstanten Monitorings und der flexiblen Bedarfsanpassung gerecht zu werden. Die inhaltliche Schwerpunktsetzung in der aktuellen Erhebung erfolgte auf den Bereich der substanzbezogenen Störungen. Erstmals seit vielen Jahren wurden die diagnostischen Kriterien nach DSM für eine Vielzahl von Substanzen (Tabak, Alkohol, Cannabis, Kokain, Amphetamine, Schmerzmittel, Schlafmittel, Beruhigungsmittel) erfasst. Dies ermöglicht eine Schätzung der Gesamtbelastung der Bevölkerung durch substanzbezogene Störungen sowie die Untersuchung von Komorbiditäten (vgl. Piontek, Kraus, Gomes de Matos & Pabst, 2013). Darüber hinaus wurde der Fragebogen so angepasst, dass neben den aktuell gültigen diagnostischen Kategorien Missbrauch und Abhängigkeit nach DSM-IV auch die zukünftig anzuwendende Kategorie substanzbezogene Störung nach DSM-5 erfasst wurde. Dies ermöglicht eine Aktualisierung und internationale Vergleichbarkeit der deutschen Ergebnisse (Agrawal, Heath & Lynskey, 2011; Chung, Martin, Maisto, Cornelius & Clark, 2012; Peer et al., 2013). Schließlich wurde auch die Inanspruchnahme von Hilfsangeboten bei Problemen mit unterschiedlichen Substanzen erfragt, sodass Einblicke in das Versorgungssystem gewonnen werden können (vgl. Gomes de Matos, Kraus, Pabst & Piontek, 2013).

Die Netto-Ausschöpfung des ESA 2012 kann insgesamt als sehr positiv bewertet werden. Besonders hervorzuheben ist, dass die Ausschöpfung über die letzten drei Erhebungen von 44.9 % im Jahr 2006 über 50.1 % im Jahr 2009 auf nunmehr 53.6 % angestiegen ist. Dieser Trend ist gegenläufig zu dem in Deutschland und anderen Ländern gefundenen teilweise dramatischen Rückgang der Responderaten in Bevölkerungsbefragungen in den letzten Jahrzehnten (Aust & Schröder, 2009; Curtin, Presser & Singer, 2005; Holbrook, Krosnick & Pfent, 2008). Insbe-

sondere der im ESA eingesetzte Methodenmix aus schriftlicher, telefonischer und Online-Befragung scheint hier positive Effekte zu haben. Ein direkter Vergleich der Beteiligungsraten mit anderen Studien ist aufgrund unterschiedlicher Studiendesigns und Grundgesamtheiten nur begrenzt möglich. Für die Studie Gesundheit in Deutschland aktuell 2010 (Robert Koch-Institut, 2012), eine telefonische Befragung zu Gesundheit und Lebenssituation unter Personen ab 18 Jahren, werden eine Reihe unterschiedlicher Ausschöpfungsquoten berichtet. So wird etwa der Anteil der realisierten Interviews im Verhältnis zu allen wahrscheinlichen Haushalten der Grundgesamtheit mit 28.9 % angegeben. Die Kooperationsraten als Anteile der realisierten Interviews an allen erreichten Haushalten bzw. Zielpersonen liegen bei 35.2 % bzw. 55.8 %. Für die Studie zur Gesundheit Erwachsener in Deutschland, die schriftliche Befragungen und medizinische Untersuchungen kombiniert, wird für die neu gezogene Stichprobe eine Responderate von 42 % angegeben (Kurth, 2012).

Die Stichprobe des ESA 2012 zeigt eine gute demographische Repräsentativität, die insbesondere durch die Gewichtung der Daten gewährleistet wird. Zwar liegt die Effektivität der Gewichtungsvariablen niedriger als in den Erhebungen zuvor, der erreichte Wert von 67.6 % ist jedoch als gut zu bewerten (Gabler et al., 1994). Die Reduktion der Effektivität im Vergleich zum ESA 2009 ist auf den Einschluss der Schulbildung als zusätzliches Anpassungsmerkmal zurückzuführen, das bisher nicht berücksichtigt wurde. Diese Änderung führt zu einer deutlich besseren Angleichung der Stichprobe an die Allgemeinbevölkerung und reduziert den sog. „Mittelschichtsbias“, also die Tatsache, dass das Bildungsniveau unter Teilnehmern von Befragungen im Vergleich zur Grundgesamtheit oftmals deutlich höher liegt (Diekmann, 2007; Kraus & Pabst, 2010). Die neue Gewichtungsvariable wird für alle Querschnittsanalysen des ESA 2012 verwendet. Für die

Trendanalysen (vgl. Kraus, Pabst, Piontek & Gomes de Matos, 2013) wurde zusätzlich ein mit früheren Erhebungen vergleichbares Redressement-Gewicht ohne Schulbildung berechnet (Effektivität 86,7 %).

Trotz der insgesamt guten Repräsentativität sind Selektionseffekte möglich, die die Prävalenzschätzungen beeinflussen können. Zum einen bestehen auch in der gewichteten Stichprobe Abweichungen von der Grundgesamtheit im Hinblick auf die Staatsbürgerschaft. Da die Befragungen ausschließlich in deutscher Sprache durchgeführt wurden, sind Personen anderer Herkunft im Survey unterrepräsentiert. Zum anderen können sich Nicht-Teilnehmer hinsichtlich ihres Konsumverhaltens systematisch von den Personen unterscheiden, die sich an der Studie beteiligt haben. Dies bestätigen die durchgeführten Non-response-Analysen. Bezüglich der Prävalenz zeigt sich, dass aktuelle Konsumenten von Alkohol und Cannabis in der Studie über-, Konsumenten von Schlafmitteln dagegen unterrepräsentiert sind. Darüber hinaus konsumieren aktuelle Tabak- und Alkoholkonsumenten, die die Teilnahme an der Studie verweigerten, (episodisch) höhere Mengen. Übereinstimmend mit anderen Untersuchungen (Lahaut et al., 2003; Lemmens, Tan & Knibbe, 1988; Zhao, Stockwell & Macdonald, 2009) legen diese Ergebnisse nahe, dass eine mangelnde Teilnahmebereitschaft an Bevölkerungsbefragungen je nach Substanz und Konsumvariable (Prävalenz oder Menge) zu Über- bzw. Unterschätzungen der wahren Werte führt. Das tatsächliche Ausmaß der dadurch entstehenden Selektionseffekte ist jedoch unbekannt, da sich stets nur ein kleiner Teil der Verweigerer an Non-response-Befragungen beteiligt und keine Informationen zu den Personen vorliegen, die überhaupt nicht mit den Studien erreicht werden.

Der Einsatz unterschiedlicher Befragungsmethoden im Epidemiologischen Suchtsurvey 2012 wirkt sich positiv auf die Ausschöpfung aus, allerdings ist es möglich, dass allein der Einsatz verschiedener Modi zu Verzerrungen führt. So zeigen sich auch nach Kontrolle soziodemographischer Variablen teils deutliche Unterschiede im Konsumverhalten. Dabei geben telefonisch und online befragte Personen insgesamt seltener Substanzkonsum an als schriftlich Befragte. Dies deutet darauf hin, dass nicht nur unterschiedliche Personengruppen bestimmte Befragungsmethoden eher nutzen, sondern dass die Verleugnungstendenz bei telefonischen Interviews und im Internet höher ist als bei schriftlichen Fragebögen. Ähnliche Modus-Effekte sind auch in internationalen Studien gut belegt (Aquilino, 1994; Beebe, McRae, Harrison, Davern & Quinlan, 2005; Curri-van, Nyman, Turner & Biener, 2004; Tourangou & Yan, 2007; Yang, Falcone & Milan, 2009). Als Erklärung für diese Befunde sind eine höhere Tendenz zu sozial erwünschten Antworten bei Anwesenheit einer dritten Person sowie ein unterschiedliches Ausmaß an wahrgenommener Vertraulichkeit und Anonymität denkbar (Aquilino, 1994; Kreuter, Presser & Tourangeau, 2008). Um Verzerrungen zu vermeiden, sollte der Befragungsmodus bei der Analyse inhaltlicher Fragestellungen als Kontrollvariable mitgeführt werden.

Zusammenfassend liegen mit dem Epidemiologischen Suchtsurvey 2012 umfangreiche bevölkerungsrepräsentative Daten zu Konsumverhalten und Störungen im Zusammenhang mit dem Gebrauch von legalen und illegalen Substanzen sowie Medikamenten vor. Der zentralen Herausforderung sinkender Ausschöpfungsraten in epidemiologischen Studien kann der Survey durch den Einsatz unterschiedlicher Befragungsmethoden standhalten. Bei der Interpretation der aus den Daten gewonnenen Ergebnisse sind jedoch mögliche Selektionseffekte aufgrund der Nicht-Berücksichtigung bestimmter Bevölkerungsgruppen sowie aufgrund von Non-response-Effekten zu beachten.

Danksagung

Der Epidemiologische Suchtsurvey 2012 wurde aus Mitteln des Bundesministeriums für Gesundheit (BMG) gefördert (AZ: IIA5–2511DSM216). Mit der Finanzierung sind keine Auflagen verbunden.

Deklaration konkurrierender Interessen

Ludwig Kraus und Daniela Piontek wurden in einem Forschungsprojekt zur Epidemiologie von Alkoholabhängigkeit von Lundbeck GmbH gefördert.

Literatur

- Agrawal, A., Heath, A. C. & Lynskey, M. T. (2011). DSM-IV to DSM-5: the impact of proposed revisions on diagnosis of alcohol use disorders. *Addiction*, 106, 1935–1943.
- Aquilino, W. S. (1994). Interview mode effects in surveys of drug and alcohol use: A field experiment. *Public Opinion Quarterly*, 58, 210–240.
- Aust, F. & Schröder, H. (2009). Sinkende Stichprobenausschöpfung in der Umfrageforschung – ein Bericht aus der Praxis. In M. Weichbold, J. Bacher & C. Wolf (Hrsg.), *Umfrageforschung – Herausforderungen und Grenzen* (S. 195–212). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Beebe, T. J., McRae, J. A., Harrison, P. A., Davern, M. E. & Quinlan, K. B. (2005). Mail surveys resulted in more reports of substance use than telephone surveys. *Journal of Clinical Epidemiology*, 58, 421–424.
- Biemer, P. & Lyberg, L. (2003). *Introduction to survey quality*. New York: John Wiley and Sons.
- Bloomfield, K. & Grittner, U. (2012, Juni). *Who harms whom with their drinking? Results from the Danish national alcohol and drug survey*. Vortrag auf dem 38th Annual Symposium of the Kettil Bruun Society for Social and Epidemiological Research on Alcohol, Stavanger.
- Bühringer, G., Augustin, R., Bergmann, E., Bloomfield, K., Funk, W., Junge, B. et al. (2000). *Alkoholkonsum und alkoholbezogene Störungen in Deutschland* (Schriftenreihe des Bundesministeriums für Gesundheit, Band 128). Baden-Baden: Nomos.
- Chung, T., Martin, C. S., Maisto, S. A., Cornelius, J. R. & Clark, D. B. (2012). Greater prevalence of proposed DSM-5 nicotine use

- disorder compared to DSM-IV nicotine dependence in treated adolescents and young adults. *Addiction*, 107, 810–818.
- Currivan, D. B., Nyman, A. L., Turner, C. F. & Biener, L. (2004). Does telephone audio computer-assisted self-interviewing improve the accuracy of prevalence estimates of youth smoking: Evidence from the UMASS Tobacco Study? *Public Opinion Quarterly*, 68, 542–564.
- Curtin, R., Presser, S. & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly*, 69, 87–98.
- Diekmann, A. (2007). *Empirische Sozialforschung: Grundlagen, Methoden, Anwendungen*. (17. Aufl.). Reinbek: Rowohlt.
- Gabler, S., Hoffmeyer-Zlotnik, J. & Krebs, D. (1994). *Gewichtung in der Umfragepraxis*. Opladen: Westdeutscher Verlag.
- Gelman, A. & Carlin, J. (2002). Poststratification and weighting adjustments. In R. M. Groves, J. L. Eltinge & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 289–203). New York: John Wiley and Sons.
- Gmel, G. & Rehm, J. (2004). Measuring alcohol consumption. *Contemporary Drug Problems*, 34, 467–540.
- Gomes de Matos, E., Kraus, L., Pabst, A. & Piontek, D. (2013). Problembewusstsein und Inanspruchnahme von Hilfe bei substanzbezogenen Problemen. *SUCHT*, 59, 355–366.
- Groves, R. M., Fowler, F. J., Couper, M. P., Lepkowski, J. M., Singer, E. & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Hoboken, New Jersey: John Wiley & Sons.
- Holbrook, A. L., Krosnick, J. A. & Pfent, A. (2008). The causes and consequences of response rates in survey by the news media and government contractor survey research firms. In J. M. Lepkowski, C. Tucker, J. M. Brick, E. De Leeuw, J. Japac, P. J. Lavrakas, M. W. Link & R. L. Sangster (Eds.), *Advances in telephone survey methodology* (pp. 499–528). New Jersey: John Wiley and Sons.
- Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S. L. et al. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine*, 32, 959–976.
- Kraus, L., Pabst, A., Piontek, D. & Gomes de Matos, E. (2013). Substanzkonsum und substanzbezogene Störungen: Trends in der deutschen Allgemeinbevölkerung 1980–2012. *SUCHT*, 59, 333–345.
- Kraus, L. & Pabst, A. (2010). Studiendesign und Methodik des Epidemiologischen Suchtsurveys 2009. *SUCHT*, 56, 315–326.
- Kreuter, F., Presser, S. & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly*, 70, 847–865.
- Kurth, B.-M. (2012). Erste Ergebnisse aus der „Studie zur Gesundheit Erwachsener in Deutschland“ (DEGS). *Bundesgesundheitsblatt – Gesundheitsforschung – Gesundheitsschutz*, 8, 980–990.
- 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., Tan, E. S. & Knibbe, R. H. (1988). Bias due to nonresponse 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.
- Peer, K., Rennert, L., Lynch, K. G., Farrer, L., Gelernter, J. & Kranzler, H. R. (2013). Prevalence of DSM-IV and DSM-5 alcohol, cocaine, opioid, and cannabis use disorders in a largely substance dependent sample. *Drug and Alcohol Dependence*, 127, 215–219.
- Piontek, D., Kraus, L., Gomes de Matos, E. & Pabst, A. (2013). Komorbide Substanzstörungen in der erwachsenen Allgemeinbevölkerung. *SUCHT*, 59, 347–354.
- Robert Koch-Institut (2012). *Daten und Fakten: Ergebnisse der Studie "Gesundheit in Deutschland aktuell 2010"*. Beiträge zur Gesundheitsberichterstattung des Bundes. Berlin: Robert Koch-Institut.
- Statistisches Bundesamt (2010). *Demographische Standards Ausgabe 2010*. (Statistik und Wissenschaft, Band 17). Wiesbaden: Statistisches Bundesamt.
- Tourangeau, R. & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133, 859–883.
- Wittchen, H.-U., Beloch, E., Garczynski, E., Holly, A., Lachner, G., Perkonig, A. et al. (1995). *Münchener Composite International Diagnostic Interview (M-CIDI, Paper-pencil 2.2, 2/95)*. München: Max-Planck-Institut für Psychiatrie, Klinisches Institut.
- World Health Organization (2000). *Guide to drug abuse epidemiology*. Geneva: World Health Organization.
- Yang, Y. M., Falcone, A. E. & Milan, L. M. (2009). Survey mode effects in two military surveys. *Section on Survey Research Methods – JSM*, 4125–4139.
- Zhao, J., Stockwell, T. & Macdonald S. (2009). Nonresponse bias in alcohol and drug population surveys. *Drug and Alcohol Review*, 28, 648–657.

Prof. Dr. Ludwig Kraus

1984 Diplom in Psychologie an der Universität Regensburg
 1992 Promotion zum Dr. phil. an der Universität Regensburg
 2006 Habilitation an der Westfälischen Wilhelms-Universität Münster
 seit 2012 Gastprofessor an der Universität Stockholm, Schweden
 seit 2012 Wissenschaftlicher Leiter des IFT Institut für Therapieforchung München

Prof. Dr. Ludwig Kraus

IFT Institut für Therapieforchung
 Parzivalstr. 25
 80804 München
 Deutschland
 Tel.: +49 (0)89–360804–30
 Fax: +49 (0)89–360804–49
 kraus@ift.de

Eingereicht: 17.05.2013
 Angenommen nach Revision: 30.09.2013