

"Politische Überzeugungssysteme Jugendlicher": einige methodologische Anmerkungen

Kirschner, Hans-Peter

Veröffentlichungsversion / Published Version

Zeitschriftenartikel / journal article

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

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Kirschner, H.-P. (1978). "Politische Überzeugungssysteme Jugendlicher": einige methodologische Anmerkungen. *ZUMA Nachrichten*, 2(2), 28-36. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-210753>

Nutzungsbedingungen:

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

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

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

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

"POLITISCHE ÜBERZEUGUNGSSYSTEME JUGENDLICHER" : EINIGE METHODOLOGISCHE ANMERKUNGEN

Im Jahre 1977 schloß ZUMA ein Projekt über "politische Überzeugungssysteme Jugendlicher" ab, bei dessen Durchführung wichtige Erfahrungen in praktischer wie auch methodologischer Hinsicht gewonnen werden konnten. Konkret ging es darum, aus der Gesamtheit aller Deutschen in der Bundesrepublik, die zum Zeitpunkt der Untersuchung zwischen 16 und 34 Jahre alt waren, eine Stichprobe zu erstellen. In dieser Stichprobe sollten Personen mit einer Schulbildung, die wenigstens den Besuch einer Fachschule einschließt, mit 50% Anteil stark überrepräsentiert sein - der Anteil dieser Gruppe an der Grundgesamtheit beträgt etwa 10%.

ZUMA hatte davon auszugehen, daß von dieser Grundgesamtheit keine Liste existiert, daß also insbesondere der disproportionale Ansatz bzgl. der Schulbildung Probleme aufwerfen würde. Trotz der zu erwartenden Schwierigkeiten operationaler und gewichtungstechnischer Art wurde beschlossen, die vorgegebene Aufgabe mit Hilfe von Zufallsstichproben anzugehen und nicht, wie in solchen Fällen vielfach üblich, eine Quotenauswahl durchzuführen. Es sollte darüber hinaus versucht werden, die nicht-stichprobenneutralen Ausfälle durch geeignet definierte "ähnliche" Einheiten aus der Gesamtheit der erfolgreich Befragten zu ersetzen, also die realisierte Stichprobe (das "Netto") nicht als eigentliche Stichprobe aufzufassen. Diese Vorgehensweise ist durchaus nicht üblich, denn sehr häufig werden die nicht-neutralen Ausfälle einfach "vergessen".

Erwartungsgemäß mußte bei der Planung und Durchführung des Projekts eine Fülle von Problemen theoretischer und praktischer Art diskutiert werden. Es stellte sich heraus, daß viele dieser Problemstellungen für ZUMA-Projekte sehr typisch sind und im übrigen im stichprobentheoretischen Sinne grundsätzlichen Charakter besitzen. Es scheint daher sinnvoll, im folgenden einige dieser stichprobentheoretischen Problemstellungen vorzustellen, zu-

mal Veröffentlichungen über Erfahrungen bei ähnlichen Projekten selten erscheinen.

Stichprobenplan und Gewichtung

Es war also eine Zufallsstichprobe aus der eingangs angegebenen Grundgesamtheit zu ziehen. Wie oben schon angedeutet, sollte die Grundgesamtheit schichtweise betrachtet werden. Außer der Variablen Schulbildung war dazu noch die Variable Alter heranzuziehen. Die Schichtenbildung kann man der folgenden Skizze entnehmen:

Schul- bildung \ Alter	16 - 30	31 - 34
	Schicht 1	Schicht 2
A	Schicht 3	Schicht 4
B		

Dabei bedeutet:

A: mindestens

Fachschulbesuch

B: bis Mittelschule

einschließlich

Für Schicht 1 und Schicht 3 waren je 900 vollständige Interviews und für Schicht 2 und Schicht 4 waren je 100 vollständige Interviews anzustreben. Dies ist ein stark disproportionaler Ansatz.

Ohne zunächst auf operationale und stichprobentheoretische Einzelheiten einzugehen, läßt sich die Stichprobenziehung wie folgt darstellen: Grundlage war die ADM-Stichprobe von 1971 (ADM=Arbeitskreis Deutscher Marktforschungsinstitute, ausführliche Behandlung in Wendt (1971)). Es wurde also ein dreistufiges Design verwendet. In der ersten Stufe wurde eine Stichprobe von 560 Stimmbezirken erstellt. Den ausgewählten Stimmbezirken wurden dann in der zweiten Stufe zufällig jeweils etwa gleichviele Haushalte entnommen. Diese Haushalte wurden kontaktiert und bei erfolgreichem Kontakt in ihrer Zusammensetzung aufgelistet. Danach wurden aus der Gesamtheit der gezogenen und vollständig erfaßten Haushalte alle diejenigen entfernt, deren

ZUMA

Mitglieder nicht zur Grundgesamtheit gehörten. Des weiteren wurde von den Haushalten, denen ausschließlich Personen mit Schulbildung B angehörten, nur jeder fünfte in den endgültigen Haushaltsdatensatz übernommen. Es ergab sich so eine Datei mit 3200 Haushalten. Aus diesen wurde dann jeweils eine Person zufällig gezogen, so daß eine Personenstichprobe der Größe 3200 resultierte, die in die oben definierten 4 Schichten zerlegt werden konnte. Innerhalb dieser Schichten wurden schließlich Teilstichproben derart erstellt, daß der vorgegebene disproportionale Ansatz verifiziert war und insgesamt 2416 Personen zur Befragung zur Verfügung standen. Schematisch läßt sich die Prozedur wie folgt verdeutlichen:

1. Stufe

Auswahl von 560 Stimmbezirken

2. Stufe

Gegeben sei ein ausgewählter Stimmbezirk: Auswahl von m Haushalten ($m \leq 35$), Befragung dieser Haushalte, Reduktion auf $m' < m$ "zulässige" Haushalte

Gegeben sei ein ausgewählter Stimmbezirk, gegeben sei die reduzierte Stichprobe der m' zulässigen Haushalte: Reduktion auf $m'' < m'$ Haushalte durch Auswahl jedes fünften Haushalts mit Personen der Grundgesamtheit ausschließlich aus Schicht 3 und Schicht 4

3. Stufe

Gegeben sei ein ausgewählter Haushalt: Auswahl einer Person der Grundgesamtheit

Gegeben sei die gesamte Personenstichprobe mit n_i Personen in Schicht i , $i=1, \dots, 4$: Auswahl von $n'_i < n_i$ Personen aus Schicht i , Befragung von rund

$$2400 = \sum_{i=1}^4 n'_i, \quad \text{Personen ("double sampling for stratification")}$$

Prinzipiell war jeder Teil dieses Designs durch einen Gewichtungsfaktor zu berücksichtigen, lieferte also jeweils eine Komponente zu einem Gewicht, das allen erfolgreich befragten Personen zugeordnet wurde, das sog. Design-Gesamtgewicht. Darin reflektiert sich das der Stichprobenziehung zugrundegelegte probabilistische Modell. Nicht erfaßt werden von diesem Modell jedoch die nicht-stichprobenneutralen Ausfälle. Wenn also - wie vorgesehen - solche Ausfälle durch "ähnliche" Einheiten aus der Menge der erfolgreich Befragten substituiert werden sollen, werden weitere Gewichte erforderlich, die sog. Substitutionsgewichte. Design-Gesamtgewicht und Substitutionsgewichte werden dann pro Fall multiplikativ zu einem Gesamtgewicht miteinander verknüpft.

Die Gewichtungskomponenten des Design-Gesamtgewichtes

Die ADM-Stichprobe von 1971 realisiert u. a. die erste Stufe eines in der Literatur seit langem eingeführten dreistufigen Stichprobenplans. Dabei wurden in der ersten Stufe Stimmbezirke (bzw. synthetisierte Stimmbezirke), die zuvor nach Regionen und Gemeindegrößenklassen angeordnet worden waren, systematisch - nach Zufallsstart - und mit einer Wahrscheinlichkeit proportional zu ihrer (geschätzten) Anzahl von Privathaushalten gezogen ("pps-Stichprobe"). Durch die vorherige Anordnung erzwingt man, daß die einzelnen "Schichten" nahezu proportional zu ihrer Größe in hinreichend umfangreichen Stichproben vertreten sind (vgl. hierzu etwa Sudman (1976), S. 134f.).

Selbstverständlich ist dies kein schichtweises Ziehen im strengen Sinne, so daß die Auswahlwahrscheinlichkeiten für die einzelnen Stimmbezirke nur Funktionen der Schrittweite I und der jeweiligen Zahl von Haushalten M sind; die Wahrscheinlichkeit für einen Stimmbezirk der "Größe" M , bei dieser Untersuchung gezogen zu werden, betrug $M/I = M/34676$. Die Wahrscheinlichkeiten M/I spiegeln also ganz allgemein ein Ziehen ohne Zurücklegen mit ungleichen Auswahlwahrscheinlichkeiten wider. Bei der Entscheidung für diese Prozedur wurde zweierlei bewußt in Kauf genommen: Die Auswahl der Einheiten erster Stufe als pps-Stichprobe ohne Zurücklegen (vgl. dazu Raj (1968), section 3.18) mußte zum einen für die anliegende Untersuchung nicht notwendig einen großen Gewinn an Präzision für die zu konstruierenden Schätzer bedeuten. Denn solche Verfahren sind besonders gut, wenn die Stimmbezirk-Totals bzgl. der untersuchten Variablen nahezu gleichmäßig proportional zum jeweiligen M sind (vgl. dazu Cochran (1972), S. 379f. und Raj (1968), S. 49 Corollary sowie Raj (1968), S. 118, Theorem 6.3). Leider war dieser Sachverhalt aus Zeitgründen und wegen fehlender Zensus-Vergleichszahlen vor Beginn der Untersuchung nicht zu verifizieren. Zum anderen wird gerade durch diese Art der Zufallsauswahl in der ersten Stufe des Designs eine Schätzung der Varianzen der Schätzer erschwert, d.h. eine Fehlerrechnung wird bereits durch die Auswahltechnik auf der ersten Erhebungsstufe kompliziert (vgl. dazu Cochran (1972), S. 308, 309, 377f.).

Da in den einzelnen Stimmbezirken aus verschiedenen Gründen nicht immer gleich viele Haushalte angelaufen werden konnten, war der rein rechnerische Vorteil einer möglichen Selbstgewichtung bzgl. der beiden ersten Stufen ebenfalls nicht zu nutzen: Ist M_i die "Größe" des Stimmbezirks i und stimmt diese mit der Größe der Auswahlgrundlage in diesem Bezirk überein, so ist bei einer zufälligen Ziehung von m_i Haushalten der Gewichtsanteil der ersten beiden Stufen für diesen Bezirk $(I/M_i) (M_i/m_i) = I/m_i$, d.h. bei gleicher Anzahl m der angelaufenen Haushalte pro Bezirk ist das Gewicht I/m für alle Bezirke dasselbe.

In dem vorliegenden Fall ergaben sich also Gewichte der Gestalt I/m_i , wenn i die Numerierung der Stimmbezirke durchläuft. Es sei jedoch darauf hingewiesen, daß diese Gewichte zwei - hoffentlich geringe - fast nicht zu behebende typische Verzerrungen aufweisen. Zum einen zeigte sich soeben, daß die "Größe" des Stimmbezirks und die Größe der Auswahlgrundlage übereinstimmen müssen, damit sich die vereinfachte Form I/m_i ergibt. Dies für die Untersuchung annehmen zu wollen, hieße aber zu unterstellen, daß die Anzahl der Haushalte pro Stimmbezirk im Jahre 1976 mit der (geschätzten) Anzahl der Haushalte pro Stimmbezirk im Jahre 1971 übereinstimmt! Die Einheiten erster Stufe wurden nämlich zu letzterem Zeitpunkt gezogen. Zum anderen erfolgte die zufällige Entnahme der Haushalte aus einer Auswahlgrundlage, in der auch Haushalte mit Personen auftraten, die z.B. sämtlich älter als 35 Jahre waren. Im Idealfalle würden diese stichprobenneutralen Ausfälle vor der Stichprobenziehung aus der Auswahlgrundlage entfernt, um die Varianz der Schätzer nicht (unkontrolliert) zu erhöhen (vgl. Raj (1968), S. 218); wären M'_i und m'_i die zu M_i und m_i korrespondierenden Zahlen nach der Bereinigung, so ergäbe sich das entsprechende Gewicht zu $(I/M_i) (M'_i/m'_i)$. Eine solche Aufbereitung des Materials konnte bei diesem Projekt u. a. aus Kosten- und Zeitgründen nicht durchgeführt werden.

Durch die folgenden Teile des Designs werden drei weitere Komponenten des Gesamtgewichtes impliziert, nämlich das "Disproportionalitätsgewicht" d , das "Haushaltsgewicht" h und das "Schichtgewicht" (n_i/n'_i) , $i=1, \dots, 4$. Dabei ist das Gewicht d gleich fünf, wenn der Befragte zu einem Haushalt mit Personen ausschließlich aus Schicht 3 und Schicht 4 gehört; sonst ist d gleich eins. Diese drei Gewichte sind weit weniger problematisch als die gerade diskutierte erste Komponente und sollen daher hier nicht näher besprochen werden.

Das Design-Gesamtwicht für eine Person P der Grundgesamtheit, die einem Stimmbezirk mit m zufällig entnommenen Haushalten entstammt und selbst in einem Haushalt mit h Personen in der Altersspanne von 16-34 Jahren lebt, ergibt sich also formal zu:

- $(I/m) \cdot d \cdot h \cdot (n_1/n_1')$, falls P zu Schicht 1 gehört, $(d=1)$
 $(I/m) \cdot d \cdot h \cdot (n_2/n_2')$, falls P zu Schicht 2 gehört, $(d=1)$
 $(I/m) \cdot d \cdot h \cdot (n_3/n_3')$, falls P zu Schicht 3 gehört und $(d=1 \text{ oder } d=5)$
 $(I/m) \cdot d \cdot h \cdot (n_4/n_4')$, falls P zu Schicht 4 gehört. $(d=1 \text{ oder } d=5)$
(Man beachte, daß die Zahlen n_i , $i=1, \dots, 4$, zufallsabhängig sind).

Die Substitutionsgewichte

In Strenge führt das Design-Gesamtgewicht nur dann zu unverzerrten Schätzern, wenn keine nicht-stichprobenneutralen Ausfälle auftreten. Dies war natürlich bei der Untersuchung nicht der Fall. Es war aber - wie gesagt - vorgegeben, die realisierte Stichprobe nicht als eigentliche Stichprobe zu behandeln, sondern auf Stimmbezirksebene und Personenebene die Ausfälle geeignet zu ersetzen (zu "matchen").

Der einfacheren Darstellungsmöglichkeit wegen sei die Technik der Ersetzung nicht-stichprobenneutraler Ausfälle von Stimmbezirken beschrieben.

Es wurden zunächst Stimmbezirksgruppen bzgl. der Variablen Bundesland, politische Gemeindegröße und Urbanisationsgrad gebildet. Alle Stimmbezirke der ursprünglichen Stichprobe lagen in genau einer dieser Gruppen. Es sei G eine Gruppe, die L nicht-stichprobenneutrale Ausfälle enthält und R Stimmbezirke, die realisiert werden konnten. Innerhalb von G wurden nun die L Ausfälle so ersetzt, als ob ihre Charakteristik (= geschätztes Stimmbezirkstotal) jeweils gleich dem arithmetischen Mittel der Charakteristiken der R "respondents" wäre, d.h. jeder der in einem dieser R Bezirke lebenden Befragten wurde bzgl. seiner Charakteristik mit dem zusätzlichen Gewicht $1 + (L/R)$ versehen.

Auf analoge Art wurden die nicht-stichprobenneutralen Ausfälle auf Personenebene ersetzt (zum Verfahren vgl. auch Konijn (1973), S. 386).

Validitätsprüfung der Stichprobe

Selbstverständlich verzerren diese Prozeduren die Schätzer. Sehr wichtig war daher eine erste Validitätsprüfung der Stichprobe mit den gewichteten Daten.

Dazu wurden durch Schätzer des Typs (Schätzer des Totals/Summe der Gewichte) die Prozentsätze der Jahrgangskohorten innerhalb der Grundgesamtheit berechnet, und zwar sowohl auf der Grundlage des Gesamtgewichts, als auch auf derjenigen des Design-Gesamtgewichts. Es zeigte sich, daß im ersten Falle die Prozentzahlen gut mit den entsprechenden Zensus-Zahlen übereinstimmten; das Weglassen der Substitutionsgewichte verschlechterte dann aber deutlich die Anpassung der Daten an den Zensus. Dies alles kann durchaus als Indiz dafür gewertet werden, daß die von ZUMA gewählte - und anderweitig oft gescheute - Strategie der konsequenten Anwendung stichprobentheoretischer Erkenntnisse bei der Erhebung und Aufbereitung des benötigten Materials zu Daten führt, die als solide Basis für nachfolgende Analysen und Schlußfolgerungen dienen können.

Schlußbemerkung

Man kann natürlich einwenden, daß eine solche Strategie wegen der meist nicht beherrschbaren Fehlerhaftigkeit der Rohdaten u. U. eine Genauigkeit vortäuscht, die realiter nicht vorhanden ist, und daß insbesondere die Ersetzung von Ausfällen die Stichprobe erst recht verzerrt. Dabei ist jedoch zu bedenken, daß ein Verzicht z. B. auf Zufallsstichproben zugleich den Verzicht auf eine fundierte Abschätzung der Varianzen der Schätzer bedeutet, daß man sich also letztlich auf sein "Glück" verläßt.

Bei der Ersetzung von Ausfällen ist ohne Frage größte Sorgfalt geboten, um nicht tatsächlich große Verfälschungen des Datenmaterials zu produzieren. Es wird in diesem Zusammenhang stets genau zu prüfen sein, ob die wichtigsten Untergruppen der Grundgesamtheit in der Menge der erfolgreich Befragten hinreichend vertreten sind. Denn offensichtlich ist die beschriebene Ersetzungstechnik inadäquat, wenn diese Bedingung nicht erfüllt ist.

ZUMA wird bei zukünftigen Untersuchungen bemüht sein, die ermutigenden Erfahrungen aus dem geschilderten Projekt weiter zu bestätigen.

Für Diskussionen und Anfragen zu den besprochenen Themenkreisen steht Hans-Peter Kirschner gerne zur Verfügung.

Literatur

- Cochran, W.G., 1972: Stichprobenverfahren, Walter de Gruyter, Berlin, New York.
- Konijn, H.S., 1973: Statistical Theory of Sample Survey Design and Analysis, North Holland Publishing Company, Amsterdam, London.
- Raj, D., 1968: Sampling Theory, Mc Graw-Hill Company, New York.
- Sudman, S., 1976: Applied Sampling, Academic Press, New York.
- Wendt, F., 1971: Darstellung eines Stichprobensystems und seiner Realisierung in der ADM-Stichprobe 1971, Hamburg.

PRETESTS BEI ZUMA

1. Zur allgemeinen Bedeutung von Pretests bei Felduntersuchungen mit

Fragebogen

In der empirischen Sozialforschung kommt dem Fragebogen als Erhebungsinstrument zur Datensammlung eine zentrale Bedeutung zu, sei es als Verfahren zur Durchführung schriftlicher oder mündlicher Umfragen, sei es als Instrument für spezifische Gruppentests und vielleicht seltener in Form des Leitfadens für den Forscher im Rahmen mehr qualitativer methodischer Ansätze. Standardisierte Techniken lassen die Anwendung von Fragebögen - und damit einen wichtigen Schritt im Forschungsprozeß - mehr und mehr problemlos erscheinen, zumal dann, wenn "bewährte", das heißt "ausprobierte" Frageinstrumente zur Verfügung stehen, auf die zurückgegriffen werden kann. Man beschränkt sich hier u. U. nur darauf, Zeitwerte zu testen und versucht, einen allgemeinen Eindruck zu gewinnen, ob der neu konstruierte Fragebogen in etwas variiertes Zusammensetzung unter den gegebenen Feldbedingungen "läuft". Je nach Qualität des Fragebogens mag damit die "Feldreife" eines Instrumentes ausreichend geprüft sein. Meist ist es jedoch erforderlich, einen differenzierter angelegten Pretest durchzuführen, um Schwächen oder sogar groben Mängeln auf die Spur zu kommen.