

Berichtssystem Weiterbildung VIII

**Durchgeführt im Auftrag des
Bundesministeriums für Bildung und Forschung**

Methodenbericht und Datendokumentation

Vorgelegt von

TNS Infratest Sozialforschung

München, Mai 2004
The/mb_za 08.30.38366

1 Vorbemerkung

Im Januar 2001 hat das Bundesministerium für Bildung und Forschung nach einer europäischen Ausschreibung Infratest Sozialforschung erneut mit der Durchführung des Forschungsprojektes „Berichtssystem Weiterbildung VIII“ (kurz "BSW") mit dem Bezugsjahr 2000 beauftragt, das seit 1979 in dreijährigem Turnus durchgeführt wird. Das Berichtssystem Weiterbildung ist derzeit die einzig verfügbare Grundlage zur Darstellung des Gesamtsystems der Weiterbildung im Bundesgebiet. Im Rahmen dieses Projektes fand die nunmehr achte bundesweit repräsentative Erhebung statt. Dabei erfolgte die Befragung zum zweiten Mal in Form von computergestützten mündlichen Interviews (CAPI). Seit dem BSW 1997 wird auch der in Deutschland lebende ausländische Teil der Bevölkerung in die Befragung einbezogen.

Das BSW-Fragenprogramm besteht aus zwei Teilen:

- einem Kernfragenprogramm, das in den verschiedenen Erhebungen im Wesentlichen unverändert abgefragt wird, um Trendanalysen zu ermöglichen;
- einem von Erhebung zu Erhebung wechselnden Fragenprogramm zu einem aktuellen Themenschwerpunkt, den im BSW 2000 das "selbstgesteuerte Lernen" bildet. Die Fragen dazu wurden im Rahmen einer Pilotstudie zu diesem Thema entwickelt (vgl. Seidel / Kuwan 2001¹). Darüber hinaus werden auch einige Fragen des letzten Themenschwerpunktes "Weiterbildungsqualität" erneut gestellt.

Der vorliegende Methoden- und Sachstandsbericht beschreibt die Anlage der Stichprobe, den Verlauf der Feldarbeit sowie die vorliegenden Datenbestand.

¹ Seidel, Sabine / Kuwan, Helmut (2001): Lernformen und Lernstrategien bei Erwachsenen. Pilotstudie zum selbstgesteuerten Lernen und den Möglichkeiten seiner statistischen Erfassung. Durchgeführt im Auftrag des Bundesministeriums für Bildung und Forschung. Hannover / München.

2 Durchführung

2.1 Pretest

Im BSW-Pretest befragten 29 Infratest-Interviewer insgesamt 58 Zielpersonen im Rahmen von computergestützten mündlichen Interviews. Dabei zeigte sich, dass der im Pretest eingesetzte BSW-Fragebogen wenig Schwierigkeiten bereitete. Dies gilt für das gesamte neue bzw. modifizierte Fragenprogramm: das Kernfragenprogramm sowie den Themenschwerpunkt selbstgesteuertes Lernen.

Insgesamt betrug die durchschnittliche Interviewdauer etwa 30 Minuten pro Interview. Somit musste entgegen den Erwartungen das Fragenprogramm für die Haupterhebung nicht mehr gekürzt werden.

2.2 Haupterhebung

Die Feldarbeiten für die bundesweite Haupterhebung begannen etwas später als beim BSW VII am 19.04.2001 und endeten am 24.07.2001. Insgesamt konnten 7.043 Interviews realisiert werden. Die Sollzahl von 7.000 Interviews wurde damit ein wenig überschritten.

Unter Verwendung des CAPI-Programms stellt sich die Interviewsituation im Wesentlichen nicht anders dar als in der herkömmlichen Befragungssituation unter Verwendung von Papierfragebögen. Bei der computergestützten Befragung können allerdings von vornherein einige Fehlermöglichkeiten bei Eingabe der Antworten ausgeschlossen werden. Die Filterführung erfolgt automatisch und die zulässigen Wertebereiche sind festgelegt, so dass ungültige Werte nicht eingegeben werden können. Ferner werden bestimmte Konsistenz- bzw. Plausibilitätsprüfungen von vornherein programmiert. Diese im Hintergrund laufenden Prüfungen geben an den entsprechenden Stellen Fehlermeldungen aus, die den Interviewer zu einer Fehlerkorrektur veranlassen.

Durch die CAPI-Programmierung verringert sich somit der Aufwand für die Datenprüfung nach Ende der Befragung. Da sich jedoch nicht alle Einzelfallprüfungen auf Kursebene vorgeben lassen, erfolgt auch bei CAPI-Interviews ein gewisser Teil der Datenprüfung nach Abschluss der Feldarbeiten.

3 Anlage der Stichprobe

Die Grundgesamtheit der bundesweiten BSW-Erhebung umfasste alle während des Befragungszeitraumes in Privathaushalten lebenden deutschsprachigen Personen der Bundesrepublik Deutschland im Alter von 19 bis unter 65 Jahren. Die Grundgesamtheit wurde mit Hilfe einer Zufallsstichprobe ermittelt. Der angestrebte Stichprobenumfang der Erhebung lag bei rd. 7.000 Fällen.

Um auch für die neuen Bundesländer differenzierte Auswertungen zu ermöglichen, wurde ein disproportionaler Stichprobenansatz gewählt, bei dem rd. 5.000 Interviews in den alten und rd. 2.000 in den neuen Bundesländern durchgeführt werden. Dadurch können repräsentative Analysen sowohl für Gesamtdeutschland als auch für die neuen und alten Bundesländer im Vergleich vorgenommen werden.

Die Auswahl der Befragungspersonen erfolgte durch eine mehrfach geschichtete, dreistufige Zufallsstichprobe. Sämtliche Gemeinden des Untersuchungsgebietes wurden nach regionalen Kriterien und ihrer Größe nach in Schichten unterteilt. Als Schichtungskriterien für die Auswahlseinheiten der ersten Auswahlstufe dienten der Reihe nach

- die Bundesländer
- die Regierungsbezirke
- die Gemeindetypen².

Die einzelnen Schichten ("Zellen") bestanden also jeweils aus allen Gemeinden desselben Typs in einem Regierungsbezirk eines Bundeslandes.

Die Aufteilung der Auswahlseinheiten der ersten Auswahlstufe (Sample-Points) erfolgte proportional zur Zahl der Haushalte in den Schichten (Zellen), und zwar so, dass auch die jeweiligen Randbesetzungen der Schichtungsmerkmale optimal den Haushaltsverteilungen der jeweiligen Randkategorien entsprechen.

²	Gemeindetyp (entsprechend BIK-Stadtregionen)
0	= 500.000 und mehr (Zone 1)
1	= 500.000 und mehr (Zone 2, 3)
2	= 100.000 bis unter 500.000 (1)
3	= 100.000 bis unter 500.000 (2, 3)
4	= 50.000 bis unter 100.000 (1)
5	= 50.000 bis unter 100.000 (2, 3)
6	= 20.000 bis unter 50.000
7	= 5.000 bis unter 20.000
8	= 2.000 bis unter 5.000
9	= unter 2.000

Zone 1 = Kernbereich der Stadtregion, Zone 2 = Verdichtungsbereich,

Zone 3 = Übergangsbereich

Die Größenzuordnung bezieht sich auf die Einwohnerzahl der betreffenden Stadtregion (Summe Zonen 1 - 3). Bei Gemeinden außerhalb dieser Regionen erfolgt die Zuordnung entsprechend der politischen Gemeindegrößenklasse (Typen 6 - 9).

Die Auswahlseinheiten der ersten Auswahlstufe bildeten die Stimmbezirke zur Bundestagswahl (alte Bundesländer) bzw. der Kommunalwahl (neue Bundesländer). Mit Wahrscheinlichkeiten proportional zur Anzahl der Haushalte in den Stimmbezirken wurde nun in jeder Zelle die erforderliche Anzahl von Sample-Points gezogen. Ein zusätzlicher Schichtungseffekt ergab sich dabei durch die Sortierung der Auswahlseinheiten nach Kreisen, Gemeinden und ggf. nach Stadtbezirken innerhalb jeder Zelle.

Insgesamt wurden 1.050 Sample-Points gezogen, davon 750 in den alten und 300 in den neuen Ländern. Dabei entfiel in den alten Bundesländern auf je ca. 39.500 Haushalte ein Sample-Point, so dass Gemeinden von dieser Größe an mit mindestens einer Auswahlseinheit in der Stichprobe vertreten sind. In den neuen Bundesländern lag dieser Wert bei etwa 22.000 Haushalten pro Sample-Point.

In der zweiten Auswahlstufe wurden durch einen Adressenvorlauf mit Zufallsauswahl die für die Stichprobe pro Sample-Point benötigten Haushaltsadressen erhoben. In diesen vorgegebenen Befragungshaushalten ermittelten die Interviewer in der dritten Auswahlstufe die zu befragenden Zielpersonen. Mit Hilfe eines systematischen Auswahlchlüssels, der für alle zum Haushalt gehörenden Personen der Grundgesamtheit die gleiche Auswahlchance gewährleistete, wurden die Zielperson eindeutig festgelegt. Jeder subjektive Einfluss der Interviewer auf die Auswahl der Befragungspersonen konnte damit ausgeschlossen werden.

Ausschöpfung der bundesweiten Repräsentativbefragung

Die Anzahl der vorgegebenen Adressen, die unterschiedlichen Ausfälle sowie die Ausschöpfung der BSW-Stichprobe kann Tabelle 1 entnommen werden.

Tabelle 1: Ausschöpfung BSW-Befragung mit dem Bezugsjahr 2000

Zielgruppe: Bevölkerung der Bundesrepublik Deutschland von 19 bis unter 65 Jahren		
	Anzahl	%
Bruttoansatz (vorgegebene Befragungsadressen)	14.638	100,0
<i>Qualitätsneutrale Ausfälle</i>		
Keine Person in der Zielgruppe	2.019	13,8
Sprachschwierigkeiten	160	1,1
Sonstige neutrale Ausfälle	340	2,3
<i>Qualitätsneutrale Ausfälle insgesamt</i>	<i>2.519</i>	<i>17,2</i>
Verbleibende Adressen	12.119	100,0
<i>Systematische Ausfälle</i>		
Im Haushalt niemanden angetroffen	1.909	15,8
Zielperson nicht angetroffen	435	3,6
Zielperson krank oder geistig nicht in der Lage, dem Interview zu folgen	101	0,8
Interview aus Zeitgründen verweigert	1.220	10,1
Interview aus anderen Gründen verweigert	1.406	11,6
Sonstige Ausfallgründe	5	0,0
<i>Systematische Ausfälle insgesamt</i>	<i>5.076</i>	<i>41,9</i>
Realisierte Interviews (Ausschöpfung)	7.043	58,1

4 Personenbezogene BSW-Daten (SPSS-Datensatz)

Das BSW-Fragenprogramm für das Bezugsjahr 2000 wurde, wie erwähnt per CAPI erhoben. Neben dem computergestützten Fragen, wurden den Befragungspersonen bei umfangreicheren Fragen Listen für den leichteren Überblick vorgelegt. Ein Ausdruck des Fragenprogramms findet sich in Anhang 1. Hieraus ist auch erkennbar, ob die Interviewer den Befragten eine Liste, die eine Kopie der Statements oder Antwortkategorien einer Frage beinhaltet, vorlegen sollten. Aufgrund dieser Listen war es nicht immer notwendig im CAPI-Programm die einzelnen Statements auszuformulieren, sondern es reichte, die auf der Liste verzeichnete Nummer im Laptop einzugeben. Dies war bei fünf BSW-Fragen der Fall (Fragen 18, 25, 26, 41 und 42). Der Vollständigkeit halber wurden daher die fünf Listen im Fragenprogramm dokumentiert.

Der Datenbestand der BSW-Daten bezieht sich auf insgesamt 7.043 Fälle. Die Daten können mit Hilfe der Variablen „PERSFAK1“ strukturgewichtet werden. Nachdem die Daten in der Stichprobe haushaltsbezogen erhoben wurden, fand in dem ersten Gewichtungsschritt eine Anpassung der Daten an die vorgegebenen Schichtungskriterien statt. Danach erfolgte der Umstieg von der Haushalts- auf die Personenebene. Hierbei werden die unterschiedlich großen Auswahlchancen von Personen in größeren bzw. kleineren Haushalten zur Beteiligung an der Befragung berücksichtigt. In die Randgewichtung der Personengewichtung einbezogen wurden die Komponenten Alter (F54), Geschlecht (F53), Bundesland (BULA) und Schulbildung (F01). Als Gewichtungsgrundlage dienten die Bevölkerungsfortschreibung und der Mikrozensus für die Bevölkerung von 19 bis 64 Jahren.

Sofern der Datensatz nicht alphabetisch sortiert aufgerufen wird, gibt er die Chronologie des Fragebogens (Anhang 1) wieder. Zusätzlich angespielte Daten, die nicht konkret erfragt wurden (z.B. Bundesland [BULA]) stehen am Ende des Datensatzes. Zusätzlich erzeugte Variablen befinden sich hinter den jeweiligen Fragen. Z. B. erscheint nach der Mehrfachnennungsfrage zu Schulabschluss die erzeugte Variable „SCHULK“ zum höchsten Schulabschluss der Probanden. Eine Übersicht über alle im Personendatensatz des BSW VIII enthaltenen Daten einschließlich ihrer Labels findet sich in Anhang 2.

Sowohl für die allgemeine als auch die berufliche Weiterbildung³ wurden die Befragungspersonen gebeten, Auskünfte über jeweils bis zu vier Kursen bzw. Veranstaltungen zu geben. Während in der Befragung pro Kurs bzw. Veranstaltung eine Schleife jeweils für die berufliche und die allgemeine Weiterbildung bis zu vier mal durchlaufen wurde, werden die Daten in der Reihenfolge des Fragenprogramms wiedergegeben. D.h., für z.B. Frage 26 stehen insgesamt vier Variablen „K126“, „K226“, „K326“ und „K426“ im Datensatz. „K1“ bis „K4“ verweist hierbei auf den ersten bis vierten Kurs der beruflichen Weiterbildung. Die verbleibenden Zeichen der Variablenbezeichnung (im Beispiel also „26“) zeigen die Nummer der Frage laut Fragebogen an. Ähnlich sind die Daten für die allgemeine Weiterbildung aufbereitet. Allerdings spricht man

³ Das Wort Weiterbildung wird in den Labels des Datensatzes häufig mit „WB“ abgekürzt.