



70

Studien und Berichte

Maïke Reimer

# Autobiografisches Gedächtnis und retrospektive Datenerhebung

## Die Rekonstruktion und Validität von Lebensverläufen



# Inhaltsverzeichnis

<b>Danksagung</b> .....	11
<b>Zusammenfassung</b> .....	13
<b>Abstract</b> .....	15
<b>Kapitel 1</b>	
<b>Befragungsmethodologie und Gedächtnispsychologie – Eine fruchtbare Allianz</b> .....	17
1.1 Gedächtnisprozesse bei der Datenerhebung .....	17
1.2 Befragungsmethodologie und Gedächtnispsychologie .....	18
1.3 Lebensgeschichtliche Ereignisdaten in retrospektiven Befragungen .....	20
1.4 Fragestellung und Vorgehen .....	23
<b>Kapitel 2</b>	
<b>Vergessen, Verzerrern, Fehldatieren – Strukturen und Prozesse autobiografischen</b>	
<b>Erinnerns</b> .....	27
2.1 Mechanismen der langfristigen Informationsspeicherung .....	27
2.2 Steuerung der Gedächtnisprozesse durch schematische Strukturen .....	29
2.3 Autobiografische Information im Langzeitspeicher .....	30
2.3.1 Repräsentation und Rekonstruktion .....	30
2.3.2 Organisationsprinzipien .....	33
2.4 Zeit und Datum im autobiografischen Gedächtnis .....	34
2.5 Fehler beim autobiografischen Erinnern: Ein Überblick über die Befundlage	37
2.5.1 Vergessensanfällige Information .....	37
2.5.2 Unterschiede zwischen Personengruppen .....	38
2.5.3 Fehler beim Datieren .....	41
2.6 Vergessen, Fehldatieren und die retrospektive Distanz .....	45
2.6.1 Vergessen und retrospektive Distanz .....	45
2.6.2 Fehldatieren und retrospektive Distanz .....	47
2.6.3 Die retrospektive Distanz und mit ihr einhergehende Faktoren .....	49
2.7 Rekonstruktionsfehler beim autobiografischen Erinnern – Ein Rahmen für Fehler in retrospektiven Befragungen? .....	51
<b>Kapitel 3</b>	
<b>Datengenerierung und Datenqualität – Konzepte, Begriffe und Befunde</b> .....	55
3.1 Daten, Realität und Fehler .....	55
3.1.1 Fehler in Ereignisdaten .....	55
3.1.2 Fehler beschreiben .....	57
3.2 Daten- und Datenfehlergenerierung in retrospektiven Befragungen .....	60
3.2.1 Die Konstruktion sozialwissenschaftlicher Daten als Variablenwerte .....	60

3.2.2	Der Gesamtbefragungsfehler und seine Komponenten . . . . .	61
3.2.3	Klassische und probabilistische Testtheorie . . . . .	63
3.2.4	Kognitive Konzepte der Befragung . . . . .	65
3.3	Ermittlung und Quantifizierung des Gedächtnisfehlers in Befragungsdaten	66
3.3.1	Reliabilität und Validität in Befragungen . . . . .	66
3.3.2	Reliabilitätsstudien: Die Wiederholungsbefragung . . . . .	70
3.3.3	Validierungsstudien: Die Registerdatenüberprüfung . . . . .	72
3.3.4	Voraussetzungen für Reliabilitäts- und Validierungsstudien mit Befragungsdaten . . . . .	74
3.4	Gedächtnisfehler in retrospektiven Ereignisdaten – Befundlage . . . . .	77
3.4.1	Fehler durch die retrospektive Rekonstruktion . . . . .	77
3.4.2	Folgen für die inhaltlichen Auswertungen . . . . .	80
3.4.3	Fehler und die Länge der retrospektiven Distanz – Befunde und grundsätzliche Überlegungen . . . . .	81
3.4.4	Validitätssteigernde Erhebungsbedingungen und -techniken . . . . .	83
3.4.5	Zusammenfassung: Gedächtnisfehler in Retrospektivbefragungen . . . . .	86
3.5	Die Deutsche Lebensverlaufsstudie (LV) . . . . .	87
3.5.1	Ziele und Konzepte der Deutschen Lebensverlaufsstudie (LV) . . . . .	87
3.5.2	Erinnerungsrelevante Erhebungsumstände in der LV . . . . .	87
<b>Kapitel 4</b>		
	„Haben Sie Kinder?“ Die Reliabilität der Angaben zu Kindern in der LV-Ost . . . . .	95
4.1	Fragestellung und Vorgehen . . . . .	95
4.2	Retrospektivbefragungen zu Kindern: Befunde und Annahmen . . . . .	96
4.2.1	Befundlage . . . . .	96
4.2.2	Datengenerierungsmodelle und Annahmen für die LV-Ost . . . . .	97
4.3	Auswertung . . . . .	102
4.3.1	Das Unterschlagen von Kindern . . . . .	102
4.3.2	Datierung von Geburtstagen und Auszügen . . . . .	106
4.3.3	Die Datierung von Auszügen und die retrospektive Distanz . . . . .	108
4.4	Diskussion . . . . .	112
<b>Kapitel 5</b>		
	„Waren Sie im Dezember 1989 erwerbstätig?“ Rekonstruktion und Reliabilität von Erwerbsverläufen im Transformationsprozess . . . . .	115
5.1	Fragestellung und Vorgehensweise . . . . .	115
5.2	Datengenerierung und Hypothesen . . . . .	116
5.2.1	Die retrospektive Erhebung von Erwerbsgeschichten in der LV-Ost . . . . .	116
5.2.2	Annahmen I: Was wird abweichend rekonstruiert? . . . . .	119
5.2.3	Annahmen II: Allgemeine und „ostspezifische“ Veränderungen . . . . .	119
5.2.4	Annahmen III: Folgen für Sequenzeigenschaften, inhaltliche Auswertungen und Edition . . . . .	121
5.3	Auswertung . . . . .	123
5.3.1	Überblick über die Daten . . . . .	123
5.3.2	Qualitative Einzelfallbetrachtungen . . . . .	124
5.3.3	Folgen für Inzidenz und Frequenz von Episoden . . . . .	130

5.3.4	Folgen für die Dauerverteilung .....	134
5.3.5	Interdependenz von Episoden und Dauerverteilungen .....	135
5.3.6	Folgen für den ersten Übergang und sein Timing .....	139
5.3.7	Folgen für die inhaltlichen Auswertungen .....	143
5.3.8	Verbesserung der Übereinstimmung durch Datenedition .....	146
5.4	Diskussion .....	146
<b>Kapitel 6</b>		
<b>Arbeitslosigkeit in der Lebensgeschichte: Eine externe Validierung mit Daten des Arbeitsamts .....</b>		
		153
6.1	Fragestellung und Vorgehen .....	153
6.2	Arbeitslosigkeit: Konzept, Definition, Erhebung .....	155
6.2.1	Arbeitslosigkeit in der LV-West .....	156
6.2.2	Arbeitslosigkeit in der LED .....	158
6.2.3	Vergleichbarkeit und Validität .....	159
6.2.4	Annahmen über Abweichungen .....	160
6.3	Auswertung I: Qualitative Einzelfallbetrachtung .....	163
6.4	Auswertung II: Konsistenz auf Monats- und Episodenebene .....	168
6.4.1	Vorgehen und Überblick über die Daten .....	168
6.4.2	Konsistenz auf Monats- und Episodenebene .....	170
6.4.3	Konsistenz und Episodenlänge .....	171
6.5	Auswertung III: Die retrospektive Distanz .....	172
6.5.1	Das Unterschlagen von Leistungsbezugszeiträumen und die retrospektive Distanz .....	172
6.5.2	Die Rekonstruktion von Arbeitslosigkeiten ohne Leistungsbezug und die retrospektive Distanz .....	176
6.6	Diskussion .....	179
<b>Kapitel 7</b>		
<b>Diskussion und Ausblick: Zusammenfassung und Implikationen .....</b>		
		185
7.1	Autobiografisches Erinnern als Fehlereinfluss in standardisierten retrospektiven Befragungen .....	185
7.2	Implikationen für die Erhebung und Edition retrospektiver Lebensverlaufsdaten .....	189
7.3	Implikationen für die Fehlerforschung .....	193
<b>Literatur .....</b>		
		197

## Tabellenverzeichnis

Tabelle 4.1:	Irrtumsrisiko, Fehlergröße und Richtung des Fehlers bei der Rekonstruktion von Kindern, ihren Geburts- und Auszugsdaten . . . . .	104
Tabelle 4.2:	Einfluss personen- und kinderseitiger Faktoren auf die Wahrscheinlichkeit, ein Kind nicht konsistent anzugeben . . . . .	105
Tabelle 4.3:	Einfluss der retrospektiven Distanz und des Alters eines Kindes beim Auszug auf das Risiko eines Auszugs, vordatiert zu werden . . . . .	112
Tabelle 5.1:	Episoden und Übergänge LV-Ost – erste und zweite Befragung . . . .	124
Tabelle 5.2:	Fehler bei Inzidenz/Frequenz der Episoden . . . . .	131
Tabelle 5.3:	Inzidenz/Frequenz der Episoden – Veränderungen zwischen erster und zweiter Befragung . . . . .	132
Tabelle 5.4:	Zusammensetzung und Komplexität der Erwerbssequenzen – Veränderungen zwischen erster und zweiter Befragung . . . . .	133
Tabelle 5.5:	Fehler bei der Dauerverteilung der drei Episodentypen . . . . .	134
Tabelle 5.6:	Äquivalent rekonstruierter Sequenzen – Ursachen und Folgen . . . .	136
Tabelle 5.7:	Interdependenz von Fehlern bei Episodenstruktur und Dauerverteilung . . . . .	137
Tabelle 5.8:	Interdependenz von Fehlern bei der Dauerverteilung . . . . .	138
Tabelle 5.9:	Erster Übergang aus Erwerbstätigkeit – Veränderungen zwischen erster und zweiter Befragung . . . . .	141
Tabelle 5.10:	Stabilität der Betriebszugehörigkeit in der ersten und zweiten Befragung . . . . .	143
Tabelle 5.11:	Übergänge aus Erwerbstätigkeit bis Dezember 1991 in der ersten und zweiten Befragung . . . . .	144
Tabelle 5.12:	Determinanten des Risikos für Erwerbsverlust in der ersten und zweiten Befragung . . . . .	145
Tabelle 5.13:	Zusammensetzung und Komplexität der Erwerbssequenzen – Übereinstimmung der unedierten und edierten Daten . . . . .	147
Tabelle 5.14:	Erster Übergang aus Erwerbstätigkeit – Übereinstimmung der unedierten und edierten Daten . . . . .	148
Tabelle 5.15:	Äquivalente Sequenzen in den unedierten und edierten Daten . . . .	148
Tabelle 6.1:	Arbeitslosigkeitsepisoden in der LV-West und der LED . . . . .	169

## Abbildungsverzeichnis

Abbildung 2.1:	Erinnern als Repräsentation und Rekonstruktion . . . . .	28
Abbildung 2.2:	Die Repräsentation autobiografischer Inhalte auf drei Ebenen und ihre Verbindungen . . . . .	31
Abbildung 2.3:	Strategien zur zeitlichen Einordnung und Rekonstruktion eines Kalenderdatums . . . . .	36
Abbildung 2.4:	Skaleneffekt bei der Rekonstruktion eines Kalenderdatums . . . . .	41
Abbildung 2.5:	Zunahme des Vergessens mit längerer retrospektiver Distanz . . . . .	46
Abbildung 2.6:	Zunahme des Datierungsfehlers mit der retrospektiven Distanz . . . . .	48
Abbildung 2.7:	Zugänglichkeit von Erinnerungen über die Lebensspanne . . . . .	50
Abbildung 3.1:	Lebensgeschichtliche Ereignisdaten (schematische Darstellung) . . . . .	57
Abbildung 3.2:	Der Prozess sozialwissenschaftlicher Datenkonstruktion . . . . .	61
Abbildung 3.3:	Die Komponenten des Gesamtbefragungsfehlers . . . . .	62
Abbildung 3.4:	Datengenerierungsmodell für Einstellungsurteile in Befragungen . . . . .	65
Abbildung 3.5:	Gedächtnisfehlerermittlung in Wiederholungsbefragungen . . . . .	71
Abbildung 3.6:	Ein Datengenerierungsmodell für prozessproduzierte Daten . . . . .	74
Abbildung 3.7:	Biografischer Kalender zur Erhebung der Lebensverlaufsdaten . . . . .	85
Abbildung 3.8:	Ein Überblick über die Erhebungen der Deutschen Lebensverlaufsstudie (LV) . . . . .	88
Abbildung 3.9:	Erhebung der Erwerbsgeschichte in der LV-West . . . . .	89
Abbildung 4.1:	Erhebung von Angaben zu Kindern in der LV-Ost . . . . .	97
Abbildung 4.2:	Zeiträume, aus denen Angaben zu Geburts- und Auszugsdaten der Kinder vorliegen . . . . .	98
Abbildung 4.3:	Datengenerierungsprozess bei der Erhebung von Kindern, ihrer Geburts- und Auszugsdaten . . . . .	99
Abbildung 4.4:	Hypothesen im Überblick . . . . .	101
Abbildung 4.5:	Verteilung von Kindschaftsverhältnis und Vitalstatus in der ersten und zweiten Befragung . . . . .	103
Abbildung 4.6:	Datierung des Geburtstags und des Auszugs – Verteilung der Abweichungen . . . . .	107
Abbildung 4.7:	Einfluss der retrospektiven Distanz auf das Irrtumsrisiko für Auszugsdaten . . . . .	108
Abbildung 4.8:	Einfluss der retrospektiven Distanz auf die Fehlergröße für Auszugsdaten . . . . .	109
Abbildung 4.9:	Einfluss der retrospektiven Distanz auf die Fehlerrichtung für Auszugsdaten . . . . .	110
Abbildung 4.10:	Einfluss des Alters eines Kindes beim Auszug auf die Fehlerrichtung . . . . .	111

Abbildung 5.1:	Das in der LV-Ost zweifach abgedeckte Zeitfenster und gedächtnisrelevante Aspekte der verstreichenden Zeit . . . . .	117
Abbildung 5.2:	Datengenerierungsprozess bei der Erhebung der Erwerbsgeschichte in der LV-Ost . . . . .	118
Abbildung 5.3:	Phasen des ostdeutschen Transformationsprozesses . . . . .	120
Abbildung 5.4:	Hypothesen im Überblick . . . . .	122
Abbildung 5.5:	Elf Arten gedächtnisverursachter Abweichung . . . . .	125
Abbildung 5.6:	Timing des ersten Übergangs aus Erwerbstätigkeit im Zeitfenster	139
Abbildung 5.7:	Datierung des ersten Übergangs aus Erwerbstätigkeit – Verteilung der Abweichungen . . . . .	140
Abbildung 5.8:	Datierung des ersten Übergangs aus Erwerbstätigkeit im Zeitfenster für übereinstimmend und abweichend benannte Übergänge .	142
Abbildung 6.1:	Vergleichszeitraum für den Vergleich von LED und LV . . . . .	154
Abbildung 6.2:	Datengenerierungsprozess bei der Erhebung von Arbeitslosigkeiten in der LV-West . . . . .	157
Abbildung 6.3:	Hypothesen im Überblick . . . . .	161
Abbildung 6.4:	Sieben Arten der Abweichung zwischen LV-West und LED . . . . .	164
Abbildung 6.5:	Übereinstimmung zwischen LV-West und LED auf Monatsebene	170
Abbildung 6.6:	Übereinstimmung zwischen LV-West und LED auf Episodenebene . . . . .	171
Abbildung 6.7:	Abhängigkeit der Übereinstimmung zwischen LV-West und LED von der Episodenlänge . . . . .	172
Abbildung 6.8:	Einfluss der retrospektiven Distanz auf die Inkonsistenzen vom Typ A (LED -> LV-West) auf Monatsebene . . . . .	173
Abbildung 6.9:	Einfluss der retrospektiven Distanz auf die Inkonsistenzen vom Typ A (LED -> LV-West) auf Episodenebene . . . . .	174
Abbildung 6.10:	Einfluss der retrospektiven Distanz auf die Inkonsistenzen vom Typ A (LED -> LV-West) auf Monatsebene getrennt nach Kohorten	175
Abbildung 6.11:	Einfluss der retrospektiven Distanz auf die Inkonsistenzen vom Typ A (LED -> LV-West) auf Episodenebene getrennt nach Kohorten . . . . .	176
Abbildung 6.12:	Einfluss der retrospektiven Distanz auf die Inkonsistenzen vom Typ B (LV-West -> LED) auf Monatsebene . . . . .	177
Abbildung 6.13:	Einfluss des Alters zum Zeitpunkt einer Arbeitslosigkeit auf die Inkonsistenzen vom Typ B (LV-West -> LED) auf Monatsebene . . .	177
Abbildung 6.14:	Einfluss des Alters zum Zeitpunkt einer Arbeitslosigkeit auf die Inkonsistenzen vom Typ B (LV-West -> LED) auf Episodenebene . .	178

## Danksagung

Ich möchte mich an dieser Stelle bei all denjenigen bedanken, die mich bei meiner Arbeit in vielfältiger Weise unterstützt haben:

- bei meinem Betreuer, Herrn Prof. Dr. Karl Ulrich Mayer, der diese Dissertation durch eine Doktorandenstelle am Max-Planck-Institut für Bildungsforschung ermöglichte und mir mit Interesse und vielfältigen Anregungen und Kommentaren zur Seite stand,
- bei Herrn Prof. Dr. Bernd Wegener, der sich freundlicherweise bereit erklärte, das Zweitgutachten zu erstellen,
- bei meinen Kolleginnen Antje Mertens, Frances McGinnity, Heike Trappe und Britta Matthes, die jederzeit mit Rat und Literaturhinweisen zur Hand waren,
- bei Ralf Künster, der eine unschätzbare Hilfe in allen Methoden- und Datenangelegenheiten war,
- bei Stefan Bender vom Institut für Arbeitsmarkt- und Berufsforschung, der die Daten der Leistungsempfängerdatei zugänglich machte und alle Nachfragen umgehend beantwortete,
- bei meiner Doktorandenkohorte, die mich aufgemuntert, abgelenkt und beraten hat: Anke Höhne, Marita Jacob, Matthias Pollmann-Schult, Justin Powell, Alessandra Rusconi, Holger Seibert und Jan S. Wagner,
- bei Anja Rampolokeng, Jochen Drewes, Ilka Holzinger, Monika Reimer und Ulrike Wehrhahn für ihren unermüdlichen Einsatz beim Korrekturlesen,
- schließlich bei Doris Gampig für die sorgfältige Erstellung der Druckvorlage.

Ihnen allen vielen Dank.





## Zusammenfassung

In dieser Dissertation untersuche ich, wie Fehler beim Erinnern autobiografischer Informationen retrospektive Ereignisdaten in standardisierten retrospektiven Lebensverlaufbefragungen beeinträchtigen. In einer solchen Befragung treffen Welten aufeinander: Um valide und aussagekräftige Analysen zu ermöglichen, müssen alle Befragten alle relevanten Episoden und Übergänge berichten und korrekt datieren. Zumindest sollten Abweichungen sich zufällig über Befragte und Antwortkategorien verteilen. Das autobiografische Erinnern ist jedoch ein selektiver und rekonstruktiver kognitiver Prozess; Information wird völlig vergessen oder verzerrt, um mit bereits gespeicherten Inhalten oder mit individuellen und normativen Schemata für Biografien übereinzustimmen.

Ich konzentriere mich auf den Prozess des Erinnerns und die Validität der Erinnerung für Episoden und Übergänge in den Lebensbereichen Kinder und Arbeitsmarktverhalten. Hierzu nutze ich die Möglichkeiten, die die *Deutsche Lebensverlaufstudie (LV)* bietet: eine Wiederholungsbefragung und eine externe Validierung von Befragtenangaben an Registerdaten.

Zunächst vergleiche ich die im Rahmen einer Wiederholungsbefragung zweimal erhobenen Angaben dreier Geburtskohorten aus Ostdeutschland zu allen ihren Kindern und zu ihrem Arbeitsmarktverhalten im Nachwendezeitraum, wobei zwischen den Befragungen etwa fünf Jahre liegen. In qualitativen Einzelfallstudien rekonstruiere ich den Prozess der Erinnerung für Episoden und Übergänge. Sowohl Episoden als auch Übergänge werden irrtümlich ausgelassen, eingefügt oder abweichend beschrieben. Übergänge werden zudem fehldatiert, wobei Datumsangaben zumeist nur wenige Monate abweichen. Quantitative Analysen belegen, dass Fehler selektiv bei denjenigen auftreten, deren Biografien viele Wechsel aufweisen und unkonventionell verlaufen. Bemerkenswert ist das Fehlen von Geschlechterdifferenzen in der Erinnerungsgüte und ein deutliches retrospektives Überberichten von Arbeitslosigkeit. Dies reflektiert die spezifischen Arbeitsmarkterfahrungen und historischen Veränderungen des Erinnerungskontextes der ostdeutschen Befragten. Modellvergleiche zeigen, dass Gedächtnisfehler die Parameterschätzung sozialwissenschaftlicher Modellrechnungen beeinflussen, dass aber eine einzelfallbasierte Datenedition sie zu einem gewissen Grade beheben kann.

Danach vergleiche ich im Rahmen einer externen Validierung die retrospektiven Angaben junger Erwachsener zu Phasen der Arbeitslosigkeit für einen Zeitraum von bis zu 20 Jahren mit den Registereinträgen der Bundesanstalt für Arbeit (BA). Als Ergebnis von Erinnerungsprozessen werden auch hier Arbeitslosigkeitsepisoden vergessen (besonders kurze und wiederholte), oder zwei oder mehr Episoden werden zu einer längeren Episode zusammengelegt. Länger zurückliegende Arbeitslosigkeit wird weniger zuverlässig erinnert; dies liegt aber nicht an der Länge des retrospektiven Intervalls, sondern am Alters- und Lebenslaufkontext beim Erleben: Arbeitslosigkeit in der Übergangphase zwischen Schule, Ausbildung und Berufseinstieg wird besonders unzuverlässig erinnert.

Die Befunde zeigen, dass besonders Befragte mit ereignisreichen, unkonventionellen Biografien mehr Unterstützung beim Erinnern im Interview selbst benötigen. Techniken des unterstützten Abrufs verbessern die Erinnerungsgüte, weil sie den Befragten helfen, die Strukturen ihrer Erinnerungsorganisation zu nutzen und den Interviewern die Möglichkeit geben, flexibel auf individuelle biografische Konstellationen einzugehen. Die Entwicklung angemessener Befragungstechniken muss auch andere Fehlerquellen – etwa die Interaktion zwischen Interviewer und Befragtem – berücksichtigen.

## Abstract

In this work, I examine how memory errors affect retrospectively collected event history data in a standardized retrospective life-course interview. Here, worlds collide: For analyses to be valid and meaningful, every respondent has to report all relevant episodes and transitions and date them correctly. At least, errors should occur at random over respondents and response categories. Autobiographical memory, however, is a selective and reconstructive cognitive process; information is forgotten completely or distorted to be consistent with information already stored and with individual or normative notions of a biography. I focus on the process of recall and the validity of recall for events and episodes in the domains of children and labor market behavior in the *German Life History Study* (GLHS).

First, I compare reports of three birth cohorts from East Germany about children and labor market behavior to the same respondents' reports five years later. In qualitative single case studies, I reconstruct the process of episode and transition recall. Respondents omit, insert, and wrongly describe both episodes and transitions; transitions are also misdated, although wrong dates are mostly only a few months off. Errors concern selectively those biographies with more transitions and less conventional episodes. Specific to the East German respondents is the absence of gender differences in recall accuracy and a marked retrospective overreporting of unemployment. This reflects East Germans' specific labor market experience and historical changes in recall context. It can be shown that recall errors can influence model estimates, but a single-case data edition counteracts them to a certain degree.

Second, I compare retrospective reports of young adults from West Germany about unemployment to official records of unemployment over a period of up to 20 years. As result of memory processes, unemployment episodes are misdated, omitted (especially short and repeated episodes), or two or more episodes are merged into one long period. While unemployment that happened longer ago tends to be remembered less accurately, this is not so much due to the retrospective interval, but rather to the age and life-course context at experience: unemployment at the transition between school, vocational training, and first employment is recalled least accurately.

These findings indicate that especially respondents with eventful unconventional biographies need more support in the interview itself. Techniques of aided recall that are based on models of autobiographical memory have been shown to greatly improve recall accuracy because they help respondents use their idiosyncratic memory pathways and permit interviewers to flexibility in dealing with individual biographical constellations and recall problems. Developing suitable techniques of aided recall must also take into account other error sources such as interviewer-respondent interaction.



# Kapitel 1

## Befragungsmethodologie und Gedächtnispsychologie – Eine fruchtbare Allianz

### 1.1 Gedächtnisprozesse bei der Datenerhebung

„Die Konstruktion von Daten ist der erste Schritt in der Konstruktion von Wissen über gesellschaftliche Verhältnisse und die gesellschaftlichen Bedingungen des sozialen Handelns der Akteure.“ (Rohwer & Pötter, 2002) Die Generierung sozialwissenschaftlicher Daten ist ein mehrstufiger Prozess, an dessen Anfang die Auswahl der sozialwissenschaftlichen Tatbestände steht, die beobachtet werden sollen. Es folgt deren Umsetzung in Variablen, die Ausarbeitung eines Erhebungsdesigns, seine Durchführung, und am Ende stehen die Einträge in die Datenbanken, die die sozialwissenschaftlichen Konzepte abbilden. Diekmann (1995) zufolge beruhen etwa 70 Prozent der zwischen 1989 und 1995 in wichtigen soziologischen Zeitschriften veröffentlichten Artikel auf Befragungsdaten<sup>1</sup>, also auf den Auskünften von Individuen über ihre Einstellungen, Meinungen und besonders ihre Erlebnisse und Verhaltensweisen. Das Antwortverhalten der befragten Menschen und die ihm zu Grunde liegenden kognitiven Prozesse bilden also eine Stufe im Datengenerierungsprozess und werden damit zu einer zentralen Quelle sozialwissenschaftlichen Wissens.

Beim Abbilden der sozialwissenschaftlichen Konzepte sind Fehler unvermeidlich und entstehen auf allen Stufen des Datengenerierungsprozesses (Groves, 1989; Lord & Novick, 1968). Somit sind das menschliche Antwortverhalten und die ihm zu Grunde liegenden kognitiven Prozesse auch eine Quelle von Fehlern und Verzerrungen in den Daten und damit im sozialwissenschaftlichen Wissen. Bei retrospektiven Befragungen, in denen Information über zurückliegende Ereignisse und Verhaltensweisen im Leben der Befragten gesammelt wird, spielen *Gedächtnisprozesse* eine wichtige Rolle bei der Generierung von Daten und Datenfehlern.

Da es keine fehlerfreie Datenerhebung gibt, geht es in der Fehlerforschung um zweierlei: erstens darum, den Fehler durch geeignete Erhebungsmethoden gering zu halten oder wenigstens dafür zu sorgen, dass er unsystematisch ausfällt; und zweitens darum, den bestehenden Fehler nach Ausmaß und Art abzuschätzen und bei der Auswertung der Daten und der Interpretation der Ergebnisse angemessen zu berücksichtigen. Gemäß dem (unter anderem) Kurt Lewin zugeschriebenen Ausspruch gibt es „nichts praktischeres als eine gute Theorie“, aus der Erwartungen, Generalisierungen und Handlungsanweisungen abgeleitet werden können. Wir benötigen also *Modelle des Datengenerierungsprozesses* bzw. partielle Modelle seiner einzelnen Stufen, die auf unserem Wissen über die daten- und datenfehlererzeugenden Mechanismen beruhen. Bei retrospektiven Befragungen, in denen die Daten unter anderem durch die Erinnerung der Befragten zu Stande kommen, muss ein solches partielles Modell auch Befunde und Annahmen über Gedächtnisprozesse in Befragungen einbeziehen.

---

<sup>1</sup> 13 Prozent verwendeten prozessproduzierte Daten und weitere 11 Prozent bezogen sich auf Texte.

## 1.2 Befragungsmethodologie und Gedächtnispsychologie

Die kognitive Gedächtnispsychologie befasst sich seit mehr als 100 Jahren mit dem Ablauf und der Güte von Erinnerungsprozessen und bietet somit ein reichhaltiges Angebot an Theorien, Konzepten und Befunden, die zur Beschreibung und Erklärung des Antwortverhaltens in sozialwissenschaftlichen Befragungen herangezogen und mit der Güte der Daten in Verbindung gebracht werden können. In den vergangenen drei Jahrzehnten hat die Befragungsmethodologie systematisch Konzepte aus der Kognitionspsychologie aufgegriffen und weiterentwickelt (Fienberg & Tanur, 1989; Jabine, Straf, Tanur, & Tourangeau, 1984; Jobe & Mingay, 1991; Jobe, Tourangeau, & Smith, 1993; Loftus, Fienberg, & Tanur, 1985; Tourangeau, Rips, & Rasinski, 2000). Dies hat die Befragungsmethodologie um bessere Erhebungsmethoden bereichert und die Kognitionspsychologie um Einsichten darüber, wie die beteiligten Prozesse grundsätzlich ablaufen. Beispielsweise hat die Verbindung von Einstellungs- und Meinungsforschern mit der psychologischen Forschung einerseits zu besseren Einstellungsfragen geführt und andererseits zu einem besseren Verständnis der Prozesse der Einstellungs- und Urteilsbildung (Schwarz & Sudman, 1996; Sudman, Bradburn, & Schwarz, 1996). Gleiches gilt für Befragungen, die auf vergangene Verhaltensweisen und Sachverhalte abzielen und dazu Konzepte und Theorien aus der Gedächtnispsychologie für sich nutzbar machen, besonders aus der Forschung zum *autobiografischen Erinnern*.

In der Gedächtnispsychologie stehen Gedächtnisfehler im Zentrum der Aufmerksamkeit: Vergessen, falsches Erinnern und die Bedingungen, unter denen es stattfindet, geben einerseits darüber Auskunft, wie viel das Gedächtnis unter welchen Bedingungen zu leisten im Stande ist, und andererseits darüber, wie Erinnern funktioniert. In der Tradition des gedächtnispsychologischen Pioniers Ebbinghaus (1885/1964) wird das Erinnern ausgiebig in experimentellen Laborstudien untersucht, in denen die Teilnehmer das später zu erinnernde Material unter kontrollierten Bedingungen erwerben und es unter ebenso kontrollierten Bedingungen wieder erinnern. Dadurch kann genau bestimmt werden, ob und wie die erinnerte von der ursprünglichen Information abweicht und unter welchen Bedingungen diese Abweichungen vermehrt, vermindert oder in anderer Form auftreten.

Die experimentellen Kontrollen schränken allerdings die Gedächtnisprozesse und -funktionen, die untersucht werden können, stark ein; insbesondere können persönlich relevante Erinnerungen und lange Behaltenszeiträume kaum betrachtet werden. Dies führte in den 1970er Jahren unter dem Stichwort „ökologische Validität“ zu einer Ausweitung der Forschung auf das Erinnern im Alltag (Gruneberg, Morris, & Sykes, 1978, 1988) und auf die Funktionen und Mechanismen autobiografischen Erinnerns (Rubin, 1986, 1996; Thompson u.a., 1998).

Für die Erinnerung an die persönliche Vergangenheit ist es allerdings meist nicht möglich, festzustellen, ob und in welcher Weise Erinnerungen von den tatsächlichen Gegebenheiten abweichen, da die tatsächlichen Gegebenheiten nicht bekannt sind. Um die Einblicke zu ergänzen, die introspektiv-phänomenologische Arbeiten in die Prozesse des autobiografischen Erinnerns gewähren (Brewer, 1986, 1996; Burt, 1992; Linton, 1986; Rubin & Kozin, 1984), wurde dennoch versucht, autobiografische Erinnerungen mit verschiedenen Methoden zu verifizieren. Manche Studien befassen sich mit der Erinnerung an öffentliche Ereignisse, deren Datum und Einzelheiten bekannt sind (Bluck, Levine, & Laulhere, 1999; Rubin & Baddeley, 1989). In anderen Studien wurden autobiografische Erinnerungen mehrfach abgefragt und verglichen (Brown & Kulik, 1977; Cohen, Conway, & Maylor, 1994; Neisser & Harsch,

1992; Wagenaar & Groeneweg, 1990) und von der Zeitstabilität auf die Korrektheit geschlossen. In anderen Studien führten die Teilnehmer Tagebücher über die jeweils wichtigen Ereignisse – teilweise über bis zu sechs Jahre –, anhand derer die Erinnerungsgüte überprüft werden konnte (Burt, Kemp, & Conbay, 2001; Loftus & Fathi, 1985; Thompson, 1982; Thompson u.a., 1998; Wagenaar, 1986). In anderen solchen Studien werden die Teilnehmer entweder nach den wichtigsten Ereignissen ihres Lebens gefragt oder nach den Ereignissen, die ihnen spontan auf ein Reizwort hin einfallen („Schlüsselwortmethode“). Man geht davon aus, dass die genannten Ereignisse die höchste Verfügbarkeit bzw. Zugänglichkeit besitzen und deshalb besonders gut erinnert werden (Crovitz & Schiffman, 1974; Fitzgerald & Lawrence, 1984; Fromholt & Larsen, 1994; Rubin, Wetzler, & Nebes, 1986). In anderen Studien wurden die Teilnehmer aufgefordert, ihre Erinnerung hinsichtlich ihrer Korrektheit selbst einzuschätzen (Brewer, 1986; Neisser & Harsch, 1992; Wagenaar, 1988).

Das autobiografische Erinnern in retrospektiven Befragungen bietet „an excellent analogue of the kinds of memory tasks in which people engage in everyday conversation“ (Jobe u.a., 1993), da:

- die retrospektive Distanz zwischen Erleben und Berichten sehr lang ist oder sein kann,
- die Beziehung zwischen dem, was die Befragten aufgefordert werden zu erinnern, und dem, was sie erinnern, nicht eindeutig ist,
- die erhobene Information meist nicht absichtsvoll gelernt und eingepägt wird,
- verschiedene Individuen verschiedene Dinge erlebt haben,
- verschiedene Individuen dieselbe Information auf verschiedenen Wegen erworben und sie unterschiedlich intensiv geübt und eingepägt haben.

Studien, die Fehler retrospektiver Befragungsdaten untersuchen, bieten also immer auch Einblicke in das autobiografische Erinnern in einer lebensechten Situation und können Laboruntersuchungen zu Vergessen und Fehlerinnern ergänzen. Es gibt mittlerweile eine ganze Anzahl von Studien, die sich mit verschiedenen Arten retrospektiver Befragungen und den in ihnen ablaufenden Erinnerungsprozessen befassen:

- (a) Wie und wie korrekt werden die *Häufigkeiten* bestimmter Ereignisse berichtet, wenn sich Befragte einer Frage gegenübersehen wie dieser: „Wie viele Arbeitstage haben Sie im vergangenen Jahr versäumt, weil Sie oder jemand anderes krank war?“ oder wenn sie Antwortkategorien wie diese verwenden sollen: „oft – selten – nie“ oder „nie – ein- bis dreimal – vier- bis zehnmal – mehr als zehnmal?“ (Blair & Burton, 1987; Menon, Raghbir, & Schwarz, 1995; Schwarz, 1990)
- (b) Wie und wie korrekt werden *einzelne Ereignisse erinnert und datiert*, wenn sich Befragte einer Frage gegenübersehen wie dieser: „Was war Ihre erste Arbeitsstelle nach Abschluss Ihrer Ausbildung?“ oder „Wann haben Sie geheiratet?“ (Belli, 1998; Friedman & Wilkins, 1985; Gibbons & Thompson, 2001; Lieury, Aiello, Lepreux, & Mellet, 1980; Linton, 1975; Loftus & Marburger, 1983; Means, Nigam, & Zarrow, 1989; Middendorf, 2000; Rubin & Baddeley, 1989; Sudman, Finn, & Lannom, 1984)
- (c) Wie und wie korrekt laufen die Erinnerungsprozesse bei der retrospektiven Erhebung von lebensgeschichtlichen Ereignisdaten ab, wenn die Teilnehmer aufgefordert werden, *lückenlose Auskünfte über die Episoden ihres Lebens in bestimmten Bereichen* zu geben? (Auriat, 1992; de Graaf & Wegener, 1989; Klein & Fischer-Kerli, 2000; Mathiowetz & Duncan, 1988; Means & Loftus, 1991; Papastefanou, 1980; Paull, 2002; Pierret, 2001; Tölke, 1980)



In dieser Arbeit werde ich mich mit den beiden letzten Befragungsformen befassen, der Erinnerung an und Datierung von einzelnen Ereignissen oder Sequenzen von Ereignissen, die zur Erhebung lebensgeschichtlicher Ereignisdaten eingesetzt werden. Ereignissequenzen bestehen zwar aus einzelnen Ereignissen, aber die Prozesse beim Erinnern einer ganzen lebensgeschichtlichen Sequenz laufen nicht einfach als wiederholter Abruf von Einzelereignissen ab, da dieses Erinnerungsformat eine lebensgeschichtliche Kontextualisierung der Einzelerinnerungen mit sich bringt (Bluck & Habermas, 2000). Auch erfordern die Eigenheiten lebensgeschichtlicher Ereignisdaten eine angemessene Beschreibung und Quantifizierung der Fehlerhaftigkeit, die ihrer zeitlich geordneten Natur Rechnung tragen. Im Hinblick auf die Erhebung von Ereignissequenzen sind die Kenntnisse über die Art und die Güte, mit der sie in standardisierten Befragungen rekonstruiert werden, noch lückenhaft und im Vergleich zur Rekonstruktion einzelner Ereignisse auch theoretisch noch weniger systematisch mit Erkenntnissen aus der Gedächtnispsychologie unterfüttert.

### 1.3 Lebensgeschichtliche Ereignisdaten in retrospektiven Befragungen

#### *Lebensgeschichtliche Ereignisdaten – Wozu und wie sie erhoben werden*

Standardisierte Befragungen sind das wichtigste Instrument zur Produktion von längsschnittlichen Individualdaten für die Lebensverlaufsforschung, da nur die Individuen selbst über die detaillierte und reichhaltige Information verfügen, die solche Analysen interessant macht. Was für Ereignisse treten im Lebenslauf auf und wie häufig? Wie lange dauern sie und welche Ereignisse folgen ihnen bzw. bleiben aus? Mit welchen Veränderungen in den gesellschaftlichen und institutionellen Rahmenbedingungen fallen sie zusammen? Wie unterscheiden sich die Lebensläufe von Männern und Frauen, Abiturienten und Hauptschülern, jüngeren und älteren Kohorten in dieser Hinsicht? Solche Informationen sind unverzichtbar für aussagekräftige Analysen von Veränderungen, Entwicklungen und Lebenswegen von Akteuren in ihren endogenen und exogenen Kontexten. Beispielsweise ist eine detaillierte Betrachtung von individuellen Ausbildungs- und Erwerbskarrieren nötig, um die Muster, Ursachen und Mechanismen von Arbeitslosigkeit über den Lebenslauf zu verstehen. Ist das Risiko, arbeitslos zu werden, hauptsächlich von feststehenden Eigenschaften von Personen und ihren Lebensverläufen abhängig, etwa ihrer sozialen Herkunft, ihrer Bildung und Ausbildung usw.? Gibt es stattdessen oder darüber hinaus eine Tendenz zur Selbstperpetuierung von Arbeitslosigkeit, in dem Sinne, dass mit häufigerer oder längerer Arbeitslosigkeit das Risiko, erneut arbeitslos zu werden, steigt bzw. die Chance, wieder in Erwerbstätigkeit zu kommen, sinkt? Da sich für solche Fragestellungen die Faktoren nicht experimentell trennen lassen, müssen reale Verläufe erfasst werden, um die zeitliche Anordnung von Episoden im Sinne einer Kausalität interpretieren zu können. Dies erfordert die Erhebung von Längsschnitt- und vorzugsweise Ereignisdaten. Dabei wird besonders auch die Erfassung der zeitlichen Geordntheit der Vergangenheit wichtig (Mayer & Huinink, 1990; Rohwer, 1994). „The multiple levels on which time enters modern data analyses adds duration, sequence and speed of events to the classical age-cohort-period model and renders cross-sectional approaches obsolete.“ (Brückner, 1995, S. 6)

Für die Erhebung lebensgeschichtlicher Ereignisdaten stehen zwei Befragungsdesigns zur Auswahl: *prospektive Panelbefragungen* und *retrospektive Einmalbefragungen* (Groves, 1989). Bei Panelbefragungen werden die Teilnehmer über die Zeit hinweg verfolgt und in angemessene-

nen Abständen wiederholt befragt. In der retrospektiven Einmalbefragung wird eine Stichprobe einmal über den interessierenden zurückliegenden Zeitraum (z.B. ein ganzes Leben) befragt<sup>2</sup>. Panelstudien haben einige Vorteile: Paneldaten stehen erst nach Jahren oder Jahrzehnten zur Verfügung und können nicht die Vorgeschichte der Gegenwart beschreiben, die Aufrechterhaltung des Kontakts mit den Teilnehmern ist aufwändig und schwierig, und durch zunehmende Stichprobenausfälle verringert sich die Zahl der Teilnehmer womöglich selektiv um bestimmte Gruppen (selektive Stichprobenattrition oder Panelmortalität). Wenn aus den Angaben der einzelnen Panelbefragungen kontinuierliche Sequenzen zusammengesetzt werden sollen, ergeben sich teilweise Schwierigkeiten, die ebenfalls die Datenqualität einschränken können (Poterba & Summers, 1984; Solga, 2001).

Retrospektive Befragungen haben vor allem zwei Nachteile: zum einen die *retrospektive Stichprobenselektivität*, da in einer in der Gegenwart gezogenen Stichprobe natürlich alle Personen fehlen, die im Verlaufe ihres Lebens durch Tod oder Wegzug aus der Population ausgeschieden sind, was selektiv bestimmte Gruppen betreffen kann. Hauptsorge allerdings ist, dass die Befragten systematisch bestimmte Ereignisse aus ihrem Leben *vergessen, falsch berichten und falsch datieren*, dass dies besonders für länger zurückliegende Zeiträume der Fall ist und dass deshalb die auf diesen Daten beruhenden deskriptiven und explanatorischen Erkenntnisse ein verzerrtes Bild der interessierenden Tatbestände zeichnen.

Dies würde sich in den beiden zentralen Gütekriterien für empirische Daten niederschlagen: Reliabilität (die Stabilität oder Zuverlässigkeit der Angaben) und Validität (die Übereinstimmung der Daten mit den tatsächlichen Sachverhalten). Lebensgeschichtliche Ereignisdaten enthalten eine große Fülle inhaltlicher und zeitlicher Informationen auf vielen verschiedenen Ebenen, die in unterschiedlichem Ausmaß und auf unterschiedliche Weise in ihrer Reliabilität oder Validität beeinträchtigt sein können. Die aussagekräftige Beschreibung und Quantifizierung von Reliabilität und Validität im Allgemeinen und für Ereignisdaten im Besonderen ist Voraussetzung dafür, solche Datenmängel mit den Erinnerungsprozessen der Befragten in Beziehung zu setzen.

#### *Lebensgeschichtliche Ereignisdaten und autobiografisches Erinnern – Zwei Welten*

In einer retrospektiven Befragung treffen zwei Welten aufeinander. Für den Sozialwissenschaftler ist es wesentlich, dass *alle* Informationen erfasst werden, die er zur Charakterisierung der zu beobachtenden Tatbestände benötigt – alle Ereignisse, alle Veränderungen –, dass inhaltliche und zeitliche Informationen so *korrekt* wie irgend möglich den tatsächlichen Gegebenheiten entsprechen oder die Abweichungen wenigstens unsystematisch ausfallen und dass die korrekte Erhebung *bei jeder Befragung jedes Teilnehmers* gewährleistet ist.

---

<sup>2</sup> Die Unterscheidung in prospektive und retrospektive Befragungen bezieht sich hier auf das *Design* der Untersuchung. Oft enthalten auch Panelbefragungen retrospektive *Fragen*, wenn auch Information über den seit der letzten Befragung verstrichenen Zeitraum erhoben wird. Diese Zeiträume sind meistens deutlich kürzer als jene, die in retrospektiven Einmalbefragungen erhoben werden, und werden deshalb allgemein für weniger problematisch gehalten. Wenn aber neue Personen zum Ausgleich für die über die Zeit verloren gegangenen Teilnehmer in die Stichprobe aufgenommen und deren Angaben über den bisher abgedeckten Panelzeitraum retrospektiv erhoben werden, treten auch in Panelbefragungen mitunter sehr lange retrospektive Perioden auf.

Das autobiografische Erinnern hingegen ist ein *selektiver, interpretativer, dynamischer* und *konstruktiver* Prozess, der aus der Fülle der sich anbietenden Erlebnisse und Informationen die relevante *auswählt, repräsentiert*, in einer Art Datenbank für den späteren Gebrauch *speichert*. Repräsentationen werden dabei so hergestellt, dass sie mit dem bereits Bekannten zusammenpassen, auch wenn dies bedeutet, dass Teile des Geschehenen ausgeblendet oder umgedeutet werden müssen. Bei Bedarf wird eine auf die aktuellen Anforderungen abgestimmte Rekonstruktion aus diesen Repräsentationen hergestellt, und zwar im Einklang damit, wie der Erinnernde jetzt die damalige Situation einschätzt (Barsalou, 1988; Bartlett, 1932; Brewer, 1996; Conway & Pleydell-Pearce, 2000; Neisser, 1986; Reiser, Black, & Kalamarides, 1986; Ross, 1991; Rubin, 1998; Thompson, Skowronski, Larsen, & Betz, 1996). Dadurch wird nicht alles, was man je erlebt, im Gedächtnis behalten, das was behalten wird, wird systematisch abweichend rekonstruiert, und die Rekonstruktionen können über mehrere Erinnerungsgelegenheiten hinweg unterschiedlich ausfallen. Dadurch wird Information, die sich besser mit bereits gespeicherter Information in Verbindung bringen lässt, besser erinnert als solche, die diesen „Voreinstellungen“ zuwider läuft (Barclay, 1986; Bluck & Habermas, 2000; Conway & Pleydell-Pearce, 2000; Lampinen, Faries, Neuschatz, & Toglia, 2000).

Als Folge dieser Prozesse wird der eigene Lebenslauf in der Rückschau durch selektives Vergessen und Umdeuten *kontinuierlicher, kohärenter* und *konventioneller* rekonstruiert, als es den tatsächlichen Gegebenheiten entspricht. Darüber hinaus gibt es Anhaltspunkte dafür, dass diese Neigung mit zunehmender zeitlicher Distanz zum ursprünglichen Ereignis immer gravierender ausfällt. Dies macht es wahrscheinlich, dass in retrospektiven Befragungsdaten spezifische, gedächtnisbezogene Fehler auftreten, indem systematisch bestimmte Ereignisse vergessen oder umgedeutet werden, und dass dies umso problematischer wird, je länger die erfragten Ereignisse zurückliegen.

#### *Empirische Zugangswege zu Gedächtnisfehlern in lebensgeschichtlichen Ereignisdaten*

Wie genau Datenfehler durch die kognitiven Prozesse entstehen, welche in den Daten enthaltene Information sie betreffen, wie verbreitet sie sind, wie groß sie ausfallen, ob sie eine systematische Verzerrung bewirken – darüber können zwei Arten von Studien Erkenntnisse liefern. *Wiederholte Befragungen* derselben Personen zu denselben Tatbeständen geben Auskunft über die Reliabilität der Angaben, und der *Abgleich der Befragungsdaten mit externen Daten*, die nicht in Befragungen gewonnen wurden, kann über deren Validität informieren.

Reliabilitäts- und Validierungsstudien werden manchmal systematisch bei der Neuentwicklung, Veränderung oder zur laufenden Qualitätskontrolle von Erhebungen durchgeführt (Auriat, 1992; Belli, Shay, & Stafford, 2001; Cannell, Fisher, & Bakker, 1965; Courgeau, 1991; de Graaf & Wegener, 1989; Freedman u.a., 1988; Papastefanou, 1980; Pierret, 2001; Tölke, 1980). Außerdem können günstige Gelegenheiten ausgenutzt werden, wenn aus anderen Gründen Personen zweimal befragt oder die Befragungsdaten mit externen Daten verknüpft wurden (Elias, 1997; Paull, 2002). In den meisten dieser Studien werden Gedächtnisprozesse als Ursache mangelnder Übereinstimmung berücksichtigt, aber in wenigen werden Hypothesen auf der Grundlage gedächtnispsychologischer Erkenntnisse entwickelt oder erinnerungsverursachte Abweichungen systematisch von anderweitig verursachten Abweichungen unterschieden (Belli, Shay, & Stafford, 1999; Klein & Fischer-Kerli, 2000; Paull, 2002; Pierret, 2001). Wo dies geschieht, werden entweder quantitative statisti-

sche Prozeduren eingesetzt (Klein & Fischer-Kerli, 2000; Paull, 2002) oder in mehr qualitativer Weise an kleinen, ausgewählten Stichproben der Datengenerierungsprozess nachverfolgt (Papastefanou, 1980; Tölke, 1980).

Das Wissen darüber, wie und wie korrekt Befragte ihre gesamten Lebensgeschichten in standardisierten Befragungen berichten, ist noch lückenhaft. Besonders für retrospektive Distanzen, die weiter als zwei oder drei Jahre in die Vergangenheit zurückreichen, und darüber, ob die Datengüte mit der verstreichenden Zeit wirklich kontinuierlich sinkt, ist nicht allzu viel bekannt. Die Annahme, dass die Länge des Zeitintervalls zwischen Erhebung und einem Ereignis mit einer höheren Fehlerhaftigkeit der Auskünfte über dieses Ereignis zusammenhängt, ist dabei weiterhin „ebenso plausibel wie ungesichert“ (Mayer, 1989).

#### 1.4 Fragestellung und Vorgehen

##### *Fragestellung und Hintergrund*

In dieser Arbeit werde ich untersuchen, ob und in welcher Weise die in der psychologischen Gedächtnisforschung nachgewiesene Neigung, den eigenen Lebenslauf in der Rückschau durch selektives Vergessen und Umdeuten zu vereinfachen und kontinuierlicher, kohärenter und konventioneller zu rekonstruieren, als es den tatsächlichen Gegebenheiten entspricht, sich auch bei der retrospektiven Erhebung sozialwissenschaftlicher Längsschnittdaten in den standardisierten Interviews der „Deutschen Lebensverlaufsstudie“ (LV) des Max-Planck-Instituts für Bildungsforschung, Berlin, bemerkbar macht und hier gedächtnisspezifische Fehler in den Daten verursacht.

Um partielle Datengenerierungsmodelle zu entwickeln, Erinnerungsprozesse und Erinnerungsgüte zu betrachten und mit der zeitlichen Stabilität von retrospektiven Befragungsdaten (Reliabilität) bzw. mit der Korrektheit (oder auch Angemessenheit) der retrospektiven Befragungsdaten (Validität) in Verbindung zu bringen, werde ich zwei einzigartige Gelegenheiten nutzen, die die LV bietet. Diese Studie erhebt seit fast 25 Jahren retrospektive Lebensgeschichten von ausgewählten Geburtskohorten in den Bereichen Wohnen, Ausbildung, Erwerbstätigkeit, Familie usw. Die Befragten rekonstruieren dabei ihr Leben innerhalb der ausgewählten Bereiche als lückenlose Abfolge von Episoden und Übergängen und berichten zu jeder Episode und jedem Übergang umfangreiche Detailinformationen (Solga, 1996; Wagner, 1989). In einer Teilstudie der LV, in der ostdeutsche Befragte Angaben zu ihren Vor- und Nachwendebiografien machten, wurde ein Teil der Befragten im Abstand von vier bis fünf Jahren zu ihren Kindern und ihrer Erwerbsbiografie zweimal befragt. Durch einen Vergleich lässt sich hier die Reliabilität der Angaben abschätzen. In einer anderen Teilstudie mit westdeutschen Befragten können die Angaben zur Arbeitslosigkeit mit externen Daten der Bundesanstalt für Arbeit (BA) auf Personenebene in Beziehung gesetzt werden; dieser Vergleich ermöglicht eine Validitätsschätzung. Folgende Fragen werden im Vordergrund stehen:

- Wie machen sich die konstruktiven interpretativen und situationsangepassten Prozesse autobiografischen Erinnerns bei der Rekonstruktion des eigenen Lebenslaufs in standardisierten Befragungen bemerkbar und wie beeinflussen sie Reliabilität und Validität?
- Welche Informationen werden zuverlässiger erinnert und welche weniger zuverlässig?
- Welche Personengruppen machen zuverlässigere und welche weniger zuverlässige Angaben?

- Wird durch eine längere zeitliche Distanz zwischen dem Ereignis und der Erinnerung der Fehler größer?
- Welchen Effekt haben Erinnerungsfehler auf die inhaltlichen Schlussfolgerungen über die beobachteten sozialwissenschaftlich relevanten Tatbestände?
- Welche Verbesserungen kann die Datenedition erzielen?
- Welche Veränderungen an der Datenerhebungsprozedur können den gedächtnisbedingten Fehlern vorbeugen helfen?

### *Aufbau der Arbeit*

Ich werde zunächst einen konzeptuellen und begrifflichen Rahmen für die Datengenerierungsmodelle entwerfen und dabei auf Literatur aus drei verschiedenen Bereichen zurückgreifen: In Kapitel 2 werde ich die relevanten Erkenntnisse über Prozesse und Güte des autobiografischen Erinnerns zusammenfassen, die einen Rahmen für die Beschreibung und Erklärung von Gedächtnisfehlern in Befragungsdaten liefern können. In Kapitel 3 werde ich relevante Konzepte aus den Theorien des Messens und Datengenerierens sowie der befragungsmethodologischen Literatur aufgreifen und innerhalb derselben die Rolle von Gedächtnisprozessen für die Datengenerierung und die Datenqualität einordnen.

Auf dieser Grundlage werde ich für drei empirische Teilstudien der LV Datengenerierungsmodelle entwerfen, die die kognitiven Prozesse bei der Beantwortung der jeweiligen Fragen beschreiben. Aus diesen werde ich Hypothesen über Ausmaß, Art und Ursache von Fehlern in den Daten ableiten und überprüfen.

In Kapitel 4 werde ich die Übereinstimmung der Angaben von vier ostdeutschen Geburtskohorten zu ihren *Kindern* untersuchen, die im Abstand von vier bis fünf Jahren zweimal erhoben wurden. Ich werde das Vergessen von Kindern und Fehler beim Datieren von Geburt und Auszug der Kinder untersuchen. Geburten und Auszüge stellen an die Datierung sehr unterschiedliche Anforderungen, was vermutlich zu sehr unterschiedlichen Fehlern führt. Außerdem werde ich prüfen, ob die Datierungsgüte für länger zurückliegende Auszüge wirklich deutlich schlechter ist als für kürzer zurückliegende Auszüge und welche Form der Zusammenhang zwischen verstreichender Zeit und Auszugsdatum hat.

In Kapitel 5 werde ich die Übereinstimmung der Angaben aus derselben Wiederholungsbefragung zu den *Erwerbsverläufen* Anfang der 1990er Jahre betrachten und untersuchen, wie sich die Angaben durch die zwischen den Befragungen verstrichene Zeit verändert haben. Für Ostdeutschland war dies ein Zeitraum rasanter und radikaler Transformation des Arbeitsmarkts und des Beschäftigungssystems, wodurch spezielle Bedingungen für die Rekonstruktion von Erwerbsbiografien gegeben sind. Infolgedessen erwarte ich, dass sich die Rekonstruktionen zwischen den beiden Befragungen auf eine charakteristische Weise verändern, die die Umwälzungen in den Erwerbsaktivitäten und deren Wahrnehmung und Einordnung widerspiegeln. Ich werde vor allem die Reliabilität der *zeitlichen* Information berücksichtigen, die in den Daten enthalten ist, nämlich die Dauerverteilungen und die Datierung von Übergängen. Neben den Ursachen und Folgen der Abweichungen werde ich außerdem untersuchen, welche Verbesserungen durch die Datenedition erzielt werden.

In Kapitel 6 werde ich mich eingehender mit der Erhebung von und Erinnerung an *Arbeitslosigkeit* beschäftigen, indem ich die Angaben zweier Kohorten aus Westdeutschland auf Personenebene mit den Daten der Bundesanstalt für Arbeit vergleiche. In der Arbeitsmarkt-

forschung ist die aussagekräftige und vergleichbare Erhebung von Arbeitslosigkeit ein wichtiges Thema, und inwieweit Befragungs- und Registerdaten abweichende Ergebnisse bringen, wird auch für den internationalen Vergleich zu Arbeitslosenquoten, -ursachen und -folgen diskutiert (McGinnity, 2001; Norwood & Tanur, 1994). Ich werde zum einen untersuchen, wie gut amtlich registrierte Perioden des Leistungsbezugs erinnert werden, und zum anderen, ob Zeiträume ohne Leistungsbezug oft dennoch als Arbeitslosigkeiten angegeben werden. Außerdem werde ich betrachten, ob die Übereinstimmung für weiter zurückliegende Episoden schlechter wird.

In Kapitel 7 werde ich die Erkenntnisse aus den drei empirischen Kapiteln zusammenfassen und ihre Implikationen für die Erhebung und Verarbeitung retrospektiver Befragungsdaten diskutieren. Dabei werde ich mich zum einen mit Methoden der Erhebung und Edition retrospektiver Daten befassen, die deren spezifischen Schwächen entgegensteuern können. Zum anderen werde ich die Methoden der Fehleranalyse, -interpretation und -dokumentation besprechen.



## Kapitel 2

### Vergessen, Verzerren, Fehldatieren – Strukturen und Prozesse autobiografischen Erinnerns

„Was war Ihre erste Arbeitsstelle nach Abschluss Ihrer Ausbildung?“ – „Wie lange haben Sie in Weilerswist gelebt?“ – „Wann haben Sie geheiratet?“ Um solche Fragen zu beantworten, muss der Befragte Information über seine eigene Vergangenheit erinnern. In diesem Kapitel werde ich zuerst beschreiben, auf welchen grundlegenden kognitiven Prozessen diese Leistung beruht, und darlegen, dass und warum die erinnerte Information erstens nicht bei jeder Erinnerung dieselbe sein muss und zweitens nicht immer der Erlebnisgrundlage entspricht. Danach werde ich darstellen, wie autobiografische Information repräsentiert und organisiert wird, und in einem gesonderten Abschnitt auf die Erinnerung an zeitliche Information wie Dauer und Datum eines Ereignisses eingehen. Auf diesem Hintergrund werde ich die Entstehung von Fehlern beim Erinnern und Datieren erklären und besprechen, in welcher Hinsicht rekonstruierte autobiografische Information sich auf charakteristische Weise von den tatsächlichen Sachverhalten unterscheidet. Ich werde darstellen, welche Ereignisse ein höheres Risiko haben, abweichend rekonstruiert zu werden, welche Personen besonders schlecht erinnern und wie erinnerte Daten von den tatsächlichen abweichen. Zuletzt werde ich diskutieren, welche Rolle die retrospektive Distanz, also der Zeitabstand zwischen Erleben und Erinnern eines Ereignisses spielt.

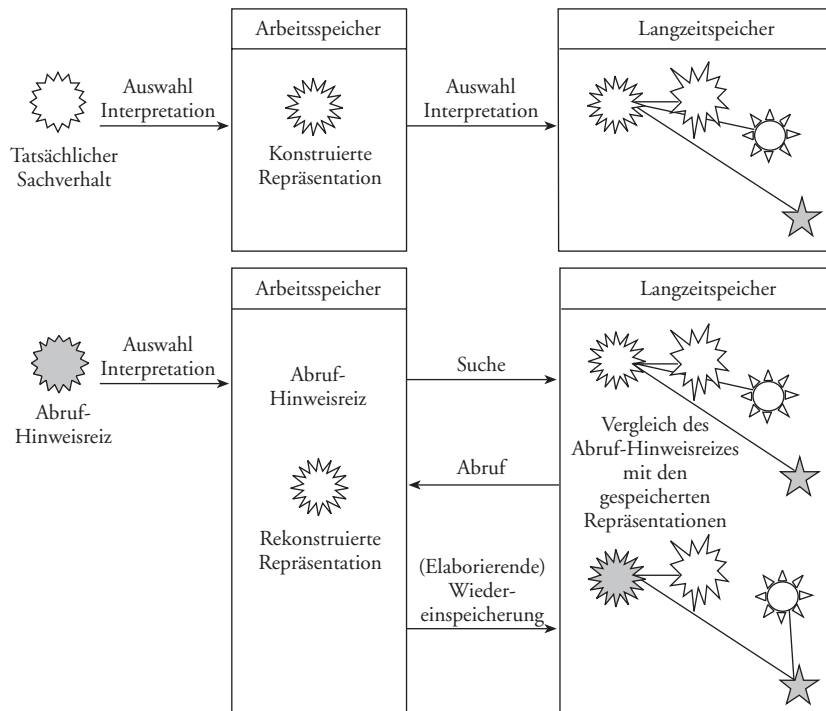
#### 2.1 Mechanismen der langfristigen Informationsspeicherung

Das langfristige Erinnern der Vergangenheit stellt eine Leistung der menschlichen Informationsverarbeitung dar, die auf zwei Komponenten beruht: zum einen auf einem Informationsspeicher, einer Art Datenbank, und zum anderen auf einem Arsenal von Prozessen, die die Information in diesem Speicher verwalten. Der Speicher besteht dabei aus zwei getrennten Teilen: einem *Arbeitspeicher*, in dem eine geringe Menge von Informationsinhalten kurzzeitig für das Bewusstsein zugänglich gehalten und zur Bewältigung alltäglicher Anforderungen genutzt werden kann (etwa die Beantwortung einer Frage, das Verstehen einer Textpassage oder das Planen einer Handlung), und einem *Langzeitspeicher* (siehe Abb. 2.1), dessen Kapazität sehr groß ist, dessen Inhalte aber nicht direkt zugänglich sind (Atkinson & Shiffrin, 1968; Gruneberg, 1976; James, 1904; Squire, 1987).

Umgekehrt werden längerfristig gespeicherte Informationen bei Bedarf von Abruf- oder Rekonstruktionsprozessen in den Repräsentationen des Langzeitspeichers gesucht, im Arbeitspeicher rekonstruiert und damit bewusst zugänglich und verwendbar. Die Rekonstruktion wird auf einen Abruf-Hinweisreiz (*retrieval cue*) hin eingeleitet, etwa einer Frage nach bestimmten autobiografischen Inhalten. Diese werden in den gespeicherten Repräsentationen entlang ihrer Verbindungen gesucht; stärkere Verbindungen werden zuerst abgesprochen, schwächere erst danach. Gelingt auf Anhieb keine befriedigende Rekonstruktion, generiert



Abbildung 2.1: Erinnern als Repräsentation und Rekonstruktion



Ein Sachverhalt wird erst im Arbeitsspeicher repräsentiert und dann in den Langzeitspeicher übertragen und mit ähnlichen Inhalten in stärkere und schwächere Verbindung gebracht. Auf einen Abruf-Hinweisreiz hin wird die passendste Repräsentation gesucht und im Arbeitsspeicher rekonstruiert. Im Anschluss wird die Information (unter Umständen in Anpassung an die neue Information verändert und mit veränderten Verbindungen) wieder in den Langzeitspeicher eingeordnet.

der Erinnernde neue internale Abruf-Hinweisreize (Atkinson & Shiffrin, 1968; Conway, 1996), die ihn hoffentlich weiterbringen. Bei der Rekonstruktion kann sowohl übergeordnete Information aus spezifischerer Information rekonstruiert werden, als auch umgekehrt spezifischere aus übergeordneter Information. „Out of a few stored bone chips we remember a dinosaur (...) Given the feeding habits of the dinosaur, we may be able to reconstruct the shapes of the bones in its foot.“ (Neisser, 1986) Häufig rekonstruierte und wieder eingespeicherte Informationen bzw. häufig abgeschrittene Rekonstruktionspfade werden „gebahnt“ und alternative Rekonstruktionspfade unterdrückt, sodass wiederholte Rekonstruktionen sich tendenziell ähneln.

Nach abgeschlossener Nutzung werden die Inhalte wieder in die Langzeitdatenbank rückübertragen. Dabei wiederholt sich prinzipiell der ursprüngliche Enkodiervorgang; wodurch sich die Repräsentation im Langzeitspeicher nach dem Abruf verändern kann: Diesen Schritt nennt man „elaborierenden Abruf“ (Ceci, Ross, & Toglia, 1987).

Eine aktuelle Rekonstruktion im Arbeitsspeicher muss dabei nicht der Erlebnisgrundlage entsprechen, und auch auf denselben Abruf-Hinweisreiz können zu unterschiedlichen Gelegenheiten unterschiedliche Rekonstruktionen hergestellt werden – zum Beispiel dann, wenn die Rekonstruktion einmal korrekt ist und einmal nicht. Abweichungen zwischen Erlebnisgrundlage und einer aktuellen Rekonstruktion bzw. zwischen zwei Rekonstruktionen zu unterschiedlichen Gelegenheiten können auf jeder Stufe des Informationsverarbeitungsprozesses entstehen: beim Enkodieren, beim Rekonstruieren und beim elaborierenden Erinnern.

Manche Autoren unterscheiden zwischen Abrufprozessen, die die Repräsentationen des Gedächtnisbestands direkt in den Arbeitsspeicher holen, und solchen, die aus den Repräsentationen die erforderliche Information rekonstruieren. Bahrck (1998) nennt dies „assoziativ-replikatives“ bzw. „inferenziell-rekonstruktives“ Erinnern. Beim replikativen Erinnern treten weniger Fehler und Verzerrungen auf, und replikatives Erinnern produziert auch zeitstabilere Rekonstruktionen. Bahrck vermutet, dass Inhalte direkt nach der Enkodierung zunächst replikativ und im Verlaufe der Zeit zunehmend rekonstruktiv erinnert werden. Durch die rekonstruktive Natur des autobiografischen Erinnerns ist erinnerte Information also immer potenziell inkorrekt und bei wiederholter Rekonstruktion potenziell instabil. Je eher die Information, die auf einen Abruf-Hinweisreiz hin rekonstruiert werden soll, den Repräsentationen entspricht, und je direkter sie aus den Repräsentationen im Langzeitspeicher in den Arbeitsspeicher geholt werden kann, desto eher wird sie korrekt und zeitstabil rekonstruiert.

## 2.2 Steuerung der Gedächtnisprozesse durch schematische Strukturen

Während die Informationsrekonstruktion also prinzipiell instabil ist und nicht notwendigerweise mit der Erlebnisgrundlage übereinstimmt, treten die Abweichungen nicht willkürlich auf, sondern folgen bestimmten Gesetzmäßigkeiten. Enkodierung und Rekonstruktion geschehen im Einklang mit der Struktur und den Organisationsprinzipien der Gedächtnisrepräsentationen untereinander: Aus den repräsentierten Erfahrungen und der Art, wie sie organisiert und strukturiert sind, werden *schematische Erwartungen* und *Vorannahmen* für die Verarbeitung neu eintreffender Information abgeleitet (Bartlett 1932; Brewer & McNamara 1984; Reiser, Black, & Abelson, 1985). Diese bestimmen, wie neue Inhalte enkodiert, repräsentiert und rekonstruiert werden.

Beim *Enkodieren* wird durch die schematischen Voreinstellungen festgelegt, welche Inhalte zur Übertragung in den Langzeitspeicher ausgewählt und wie sie repräsentiert werden, mit welchen anderen, bereits repräsentierten Inhalten der neue Inhalt verknüpft wird, wie diese Verbindungen aussehen und welcher Art diese Verbindungen sein sollen. Eine nichtbestandene Prüfung kann als schweres selbstverschuldetes Versagen interpretiert und in enge kausale Beziehung mit nachfolgenden beruflichen Rückschlägen gebracht werden, wenn die schematischen Voreinstellungen der Person eine solche Verbindung nahe legen.

Auch die *Rekonstruktion* wird von den schematischen Strukturen geleitet. Diese bestimmen, wo eine Information gesucht wird; alle anderen Erinnerungen werden von vornherein ausgeschlossen und der Suchraum so eingeengt. Wer sich an ein besonderes Versagenserlebnis erinnern möchte, wird die oben erwähnte nicht bestandene Prüfung nur rekonstruieren, wenn er sie auch als solches repräsentiert hat. Wer sich an die Ursache seines beruflichen Scheiterns

erinnern möchte, wird die nicht bestandene Prüfung dann erinnern, wenn kausale Verbindungen zwischen nicht bestandenen Prüfungen und beruflichem Scheitern bestehen.

Die Organisationsstrukturen sind in einem gewissen Grade veränderlich in Anpassung an neue Information. Wenn der Erinnernde feststellt, dass seine Kommilitonen, die die Prüfung bestanden haben, ebenfalls holprige Karrieren haben, mag er die schematischen Verbindungen zwischen nicht bestandener Prüfung und beruflichen Rückschlägen lockern und die Annahme modifizieren, dass berufliches Scheitern auf Scheitern bei Prüfungen folgt. Da aber Information tendenziell in Einklang mit den Vorannahmen verarbeitet wird, haben diese eine starke Neigung zur Verfestigung und dies hat auch Folgen für die Richtigkeit und zeitliche Konstanz von rekonstruierten Erinnerungen:

- Informationen, die nicht zu den Erwartungen und Vorannahmen der Struktur passen, werden mit erhöhter Wahrscheinlichkeit *gar nicht im Langzeitspeicher repräsentiert* und sind *schwerer rekonstruierbar* als schemakonsistente.
- Informationen, die nicht zu den Erwartungen und Vorannahmen der Struktur passen, werden sowohl beim Enkodieren als auch bei der Rekonstruktion *in Richtung des Schemas verzerrt*, das heißt durch Akzentuieren schemakompatibler Aspekte eines Inhalts und Herunterspielen schemainkompatibler Aspekte „passend“ oder „sinnvoll“ gemacht.
- Wenn die gesuchten Inhalte nicht oder nicht exakt genug abrufbar sind, werden beim Rekonstruieren oder elaborierenden Wiedereinspeichern *Schätzungen in Übereinstimmung mit den schematischen Erwartungen* angestellt.

Durch die schematische Informationsverarbeitung wird Information so erinnert, dass sie zur aktuellen Einschätzung des Erinnernden über die damalige Situation passt (Barsalou, 1988; Bartlett, 1932; Brewer, 1996; Conway & Pleydell-Pearce, 2000; Neisser, 1986; Reiser, Black, & Kalamaridis, 1986; Ross, 1991; Rubin, 1998; Thompson, Skowronski, Larsen, & Betz, 1995). Dies läuft auf zweierlei hinaus: Erstens haben Rekonstruktionen, die nicht den aktuellen schematischen Erwartungen entsprechen, ein höheres Risiko, von der Erlebnisgrundlage abzuweichen oder bei wiederholter Rekonstruktion auf unterschiedliche Weise rekonstruiert zu werden. Zweitens werden Rekonstruktionen eher so von der Erlebnisgrundlage abweichen, dass sie besser zu den aktuellen organisierenden Schemata passen als die Erlebnisgrundlage.

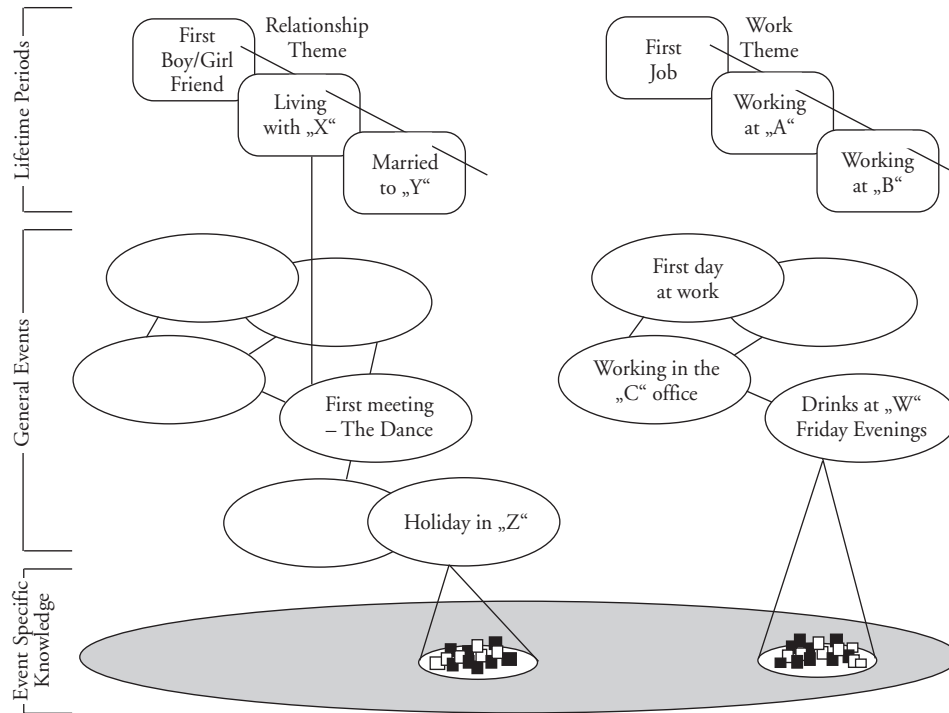
In den folgenden Abschnitten werde ich die Natur der Repräsentationen im autobiografischen Gedächtnis, ihre schematischen Organisationsstrukturen und das zeitliche Einordnen von Erinnerungen beschreiben. Vor diesem Hintergrund werde ich Art und Entstehung von Fehlern beim Erinnern besprechen und diskutieren, welche Rolle die Länge der Zeit zwischen Erleben und Erinnern eines Ereignisses spielt.

## 2.3 Autobiografische Information im Langzeitspeicher

### 2.3.1 Repräsentation und Rekonstruktion

Unter autobiografischem Gedächtnis wird die langfristige Speicherung *aller* Informationen, die sich auf den Erinnernden selbst beziehen, verstanden. Tulving (1972, 1985, 1989) unterscheidet im Langzeitspeicher zwei Arten von Informationen, deren Speicherung und Rekonstruktion unterschiedlich ablaufen: *episodische* und *semantische* Information. Im episodischen Gedächtnis werden selbst erlebte Erinnerungen abgespeichert, beispielsweise die Erinnerung

Abbildung 2.2: Die Repräsentation autobiografischer Inhalte auf drei Ebenen und ihre Verbindungen



Quelle: Conway, 1996.

an den letzten Urlaub oder den ersten Schultag. Im semantischen Gedächtnis hingegen befinden sich Fakten und Wissen, das sich nicht auf konkrete Erlebnisse bezieht, wie zum Beispiel Schulnoten, Hausnummern und Kalenderdaten<sup>1</sup>.

Zum autobiografischen Erinnern gehört also die Erinnerung an selbst Erlebtes und an Fakten ohne Erlebnisgrundlage. Implizit oder explizit beschränken sich allerdings viele Autoren in ihren Arbeiten auf selbst erlebte Inhalte und klammern Fakten aus. So nennt Rubin (1996) folgende charakteristische Merkmale für autobiografische Erinnerungen: Sie besitzen eine bildhafte Komponente und spezifische Details, an ihre Wahrhaftigkeit wird fest geglaubt, das Erinnern geht mit emotionalen Begleiterscheinungen und einem Empfinden des Wiedererlebens einher. Brewer unterscheidet ein „*Erinnerungsgedächtnis*“ (*recollective memory*) für selbst erlebte Inhalte und ein *Faktengedächtnis* für autobiografische Information ohne Erleb-

<sup>1</sup> Außerdem nimmt Tulving einen Speicher für prozedurale Inhalte an, in dem motorische und andere sinnnahe Kenntnisse und Fertigkeiten wie Gehen, Fahrradfahren, Gerüche, Geschmäcke gespeichert werden. Da derartige Informationen selten zum autobiografischen Erinnern gerechnet oder in Befragungen erhoben werden, werde ich dies hier nicht berücksichtigen.

nisgrundlage (Brewer, 1986, 1996), beschäftigt sich aber dann hauptsächlich mit den Inhalten des Erinnerungsgedächtnisses.

Übereinstimmend finden zahlreiche Autoren, dass autobiografische Informationen in der Form von untereinander in Beziehung stehenden *Episoden* organisiert und repräsentiert sind. Brewer nimmt an, dass Episoden in zweierlei Form auftreten: als *persönliche Erinnerungen* und als *generische Erinnerungen*. Persönliche Erinnerungen sind einzelne selbst erlebte Episoden, die reich an Kontexten und Details aller Sinnesmodi sind. Ihre Rekonstruktion bringt ein Empfinden von Wiedererleben mit sich und sie werden als eindeutig raumzeitlich verortet und umgrenzt wahrgenommen, selbst wenn ihnen keine Kalenderdaten zugewiesen werden können. Beispiele für persönliche Erinnerungen wären „mein erster Arbeitstag“ oder „meine Reise nach Kanada“. Generische Erinnerungen hingegen bilden das Allgemeine, Typische von mehrfach in ähnlicher Form aufgetretenen Ereignissen ab – etwa „Weihnachten mit der Familie“ oder „Vorstellungsgespräche“. Innerhalb der generischen Repräsentation verlieren die Einzelerinnerungen an Distinktheit und können nur noch schwer voneinander differenziert und eindeutig raumzeitlich verortet werden. Neisser (1986) bezeichnet derartige Repräsentationen auch treffend als „Repisoden“.

Conway und seine Mitarbeiter (Conway, 1996; Conway & Bekerian, 1987; Conway & Pleydell-Pearce, 2000) finden in der autobiografischen Informationsbasis episodische Repräsentationen auf drei Ebenen der Spezifität: *Lebensabschnitte*, *allgemeine Ereignisse* und *ereignisspezifische Details* (siehe Abb. 2.2)<sup>2</sup>. Lebensabschnitte sind am allgemeinsten und zeitlich am ausgedehntesten. Sie enthalten inhaltliche und zeitliche Informationen über andere Menschen, Orte, Aktivitäten, Pläne und Ziele, die einen bestimmten Zeitabschnitt charakterisieren – etwa „als ich in Firma A. arbeitete“ und „als ich in Firma B. arbeitete“. Sie bilden sozusagen „Kapitel“ eines Lebens, unter die bestimmte Erinnerungen „fallen“ können. Lebensabschnitte verlaufen innerhalb eines thematischen Bereichs wie „Arbeit“ oder „Beziehungsleben“ und können sich deshalb zeitlich überlappen. Die Grenzen von solchen Lebensabschnitten werden von Übergängen in dem betreffenden Bereich gekennzeichnet, wenn die ihn definierenden anderen Personen, Aktivitäten, sozialen Rollen, Pläne und Ziele sich so verändern, dass ein neuer Lebensabschnitt beginnt, wobei diese Übergänge auch unscharf und verschwommen (*fuzzy*) sein können.

Eine Ebene spezifischer treten autobiografische Informationen (ähnlich wie in Brewers Erinnerungsgedächtnis) einmal als einzelne genau umgrenzte Episoden, zum anderen wiederholte Ereignisse in generischer Repräsentation auf. Auf der spezifischsten Ebene werden die autobiografischen Eindrücke und Details als „ereignisspezifisches Wissen“ in einheitlicher Form kontextfrei und wahrnehmungsnah in einem undifferenzierten „Datenpool“ repräsentiert.

Die Ereignisse, Episoden und Lebensabschnitte, als die autobiografische Information repräsentiert ist, sind *subjektive Konstruktionen*.

„An event need not be a continuous temporal unit (...) many events are intrinsically discontinuous, they are no less real on that account. It is frequently suggested that events like these have no objective reality – that they are brought into existence only by the way we perceive and talk about them.“ (Neisser, 1986)

<sup>2</sup> Conway und Pleydell-Pearce (2000) zitieren neurologische Fallstudien als Beleg dafür, dass diese drei Ebenen der Spezifität sogar getrennte neurologische Grundlagen haben. Personen mit bestimmten Hirnläsionen entwickelten spezifische Amnesien für Information auf einer oder zwei der drei Ebenen, während die Rekonstruktion von Episoden auf den anderen Ebenen weiterhin möglich war.

Die Bestandteile des autobiografischen Informationsbestands – Details, Episoden und Lebensabschnitte – verweisen durch ihre Verknüpfungen auf andere Einheiten derselben oder einer anderen Ebene. Beispielsweise sind ausgehend von einem Lebensabschnitt all die zentralen Aktivitäten und Episoden erreichbar, die ihn definieren und von anderen Lebensabschnitten abgrenzen. Von allgemeinen Ereignissen ausgehend werden die dazugehörigen Ereignisdetails auf dem spezifischsten Level zugänglich. Je nach der Art des Abruf-Hinweisreizes kann der Suchprozess auf jeder der drei Ebenen begonnen werden; Abruf-Hinweisreize, die Lebensabschnitte aktivieren, sind dabei am effektivsten (Conway & Bekerian, 1987). Diese Verbindungen verlaufen entlang hierarchisch-thematischer, temporaler und horizontaler zwischenbereichlicher Rekonstruktionspfade (Barsalou, 1988; Belli, 1998; Conway, 1996; Conway & Pleydell-Pearce, 2000). Neisser (1986) geht davon aus, dass die Episoden in einer verschachtelten (*nested*) Struktur angeordnet sind, wobei kleinere Einheiten in größere eingebettet werden. Hierarchieebenen können dabei aber auch überlappen und Ereignisse mehreren übergeordneten Einheiten zugehören. So können je nach Situationsanforderung unterschiedliche Erinnerungen rekonstruiert werden und jeweils andere Details aus dem Datenpool aktiviert oder Lebensabschnitte anders begrenzt werden. Auf diesen Wegen werden in jeder Erinnerungssituation die erforderlichen Rekonstruktionen im Arbeitsspeicher aufs Neue konstruiert.

Bei der Rekonstruktion autobiografischer Information spielen also die Verbindungen von Repräsentationen untereinander eine entscheidende Rolle. Nach welchen übergeordneten Gesichtspunkten werden Episoden miteinander in Verbindung gebracht und aufeinander bezogen? In der Literatur finden sich bei verschiedenen Autoren zwei übergeordnete Organisationsprinzipien für den autobiografischen Erinnerungsbestand: zum einen das *Selbstschema*, zum anderen *narrative Schemata*. Im nächsten Abschnitt werde ich diese zwei zentralen Ordnungsprinzipien beschreiben und danach gesondert auf die Repräsentation und Rekonstruktion zeitlicher Information im autobiografischen Gedächtnis eingehen.

### 2.3.2 Organisationsprinzipien

Da es sich bei autobiografischer Information um selbstbezogene Informationen handelt und die Erinnerung an die eigene Vergangenheit einen wichtigen Aspekt der Selbstdefinition und Identitätsbildung darstellt, werden die Repräsentationen durch die dynamischen, veränderlichen *Selbstschemata* organisiert, die zum Zeitpunkt des ersten oder erneuten Enkodierens und zum Zeitpunkt der Rekonstruktion aktiv sind. Dabei handelt es sich um die generalisierten Erwartungen und Vorstellungen, die eine Person über sich selbst, ihre Eigenschaften, Rollen, Fähigkeiten und Ziele im Allgemeinen und in spezifischen Kontexten (etwa beim Sport, bei der Arbeit oder in der Familie) entwickelt hat (Barsalou, 1988; Conway & Pleydell-Pearce, 2000; Habermas & Bluck, 2000; Markus, 1977; Singer & Salovey, 1993). Die Organisation autobiografischer Repräsentationen entlang der für das Individuum wichtigen Selbstaspekte findet sich schon auf der Ebene allgemeiner Episoden: Hier werden thematische und zeitliche „Cluster“ von Erinnerungen gebildet, die nach selbstrelevanten Aspekten gruppiert sind (Brown & Schopflocher, 1998). Besonders wichtig sind dabei die aktuellen Ziele und Pläne (Barsalou, 1988). Robinson (1992) beschreibt, dass diese lokalen Ordnungen sich oft zu *mini-histories* von Zielerreichung oder -nichterreichung um wichtige zentrale Episoden herum organisieren. Eine besondere Rolle schreibt er dabei *first time events* zu, die als elaborierte

Minigeschichten repräsentiert werden: der erste Kuss, das erste Mal ein Auto alleine fahren usw., da solche „ersten Ereignisse“ in der Selbstdefinition eine besondere Rolle spielen.

Mit dem Selbstaspekt eng verwoben ist der narrative Aspekt der Gedächtnisorganisation, da ein kohärentes Lebensnarrativ eine zentrale Funktion des Selbst darstellt (Bluck & Habermas, 2000; Neisser & Fivush, 1994; Robinson & Taylor, 1998). Man erinnert eben nicht nur das zeitliche Aufeinanderfolgen von Lebensabschnitten oder Ereignissen, sondern an die *Geschichte* seines Lebens, in der die einzelnen Bestandteile begründet und dramaturgisch als Vorgeschichte zur Gegenwart arrangiert werden müssen: „One’s autobiography is not simply the sum of one’s memories. (...) Autobiographies (...) are versions of a life that employ cultural forms to present a persuasive und coherent account of a life.“ (Robinson & Taylor, 1998, S. 126) Bluck und Habermas nehmen an, dass ein *life story*-Schema die Organisation autobiografischer Erinnerungen zu einer Lebensgeschichte vornimmt und die Vergangenheit in einen konsistenten Zusammenhang mit der Gegenwart stellt (Bluck & Habermas, 2000, 2001). Das *life story*-Schema enthält auch kulturspezifische Normen bzw. Erwartungen darüber, welche Inhalte in eine Biografie gehören und welche nicht (Conway & Bekerian, 1987) und welche Erfahrungen in welcher Reihenfolge und zu welchen Zeitpunkten gemacht werden (Heckhausen, 1993; Kalicki, 1996; Settersten & Mayer, 1997). Narrative Strukturen stellen für autobiografische Erinnerungsinhalte Konsistenz, Plausibilität und Sinnhaftigkeit innerhalb eines kulturell akzeptierten Systems, wie eine Geschichte im Allgemeinen und eine Biografiegeschichte im Besonderen erzählt wird.

Die Organisation autobiografischer Information auf der Grundlage von Selbstschemata und narrativen Schemata hat zur Folge, dass selektiv das besser erinnert wird, was zum Zeitpunkt des aktuellen Abrufs für die Ziele und Pläne des jeweils aktuellen Selbst(aspekts) im Vordergrund steht. Außerdem werden selektiv jene Ereignisse besser erinnert, die in die subjektive Dramaturgie der Erzählung passen, und Information, die nicht passt, wird so uminterpretiert, dass sie sich besser in die Geschichte einfügt. Wenn sich bei zwei unterschiedlichen Gelegenheiten die Selbstaspekte, die die Rekonstruktion anleiten, unterscheiden, können sich dadurch auch die Rekonstruktionen voneinander unterscheiden. Zwar sind die schematischen Strukturen wie beschrieben tendenziell sehr stabil und stabilisieren sich auch selbst; sie sind aber nicht unveränderlich, und vor allem können in unterschiedlichen Situationen unterschiedliche Aspekte im Vordergrund stehen (Markus, 1977).

Als Folge der selektiven Entfernung und Umdeutung von Information, die nicht in die aktuelle Entwicklungslogik und die aktuellen Selbstschemata passt, sind rekonstruierte Autobiografien im Vergleich zum tatsächlich Geschehenen folgerichtiger, konsistenter mit der Gegenwart und stärker im Einklang mit subjektiven und normativen Entwicklungsgeschichten; es kommt zur Vereinfachung, Konventionalisierung und Glättung der Lebensverläufe (Barsalou, 1988; Middendorf, 2000; Neisser & Fivush, 1994).

#### 2.4 Zeit und Datum im autobiografischen Gedächtnis

„Nothing appears more fundamental to a biography than time“, stellen Larsen, Thompson und Hansen (1996) fest. Dennoch ist der Zeitpunkt, zu dem ein Ereignis stattgefunden hat, das Detail, das am schlechtesten erinnert wird, und das Datum oder der Wochentag ein sehr schlechter Abruf-Hinweisreiz für die Rekonstruktion eines Ereignisses (Barsalou, 1988; Brewer,

1988; Wagenaar, 1986). Dies legt nahe, dass die zeitliche Dimension des autobiografischen Gedächtnisses nicht in direktem Bezug zu einem allgemeinverbindlichen Kalender im Gedächtnisspeicher abgebildet ist. Nur wenige Erinnerungsepisoden sind mit einer Art „Zeitetikett“ in Form eines Kalenderdatums oder einer Zeitdauer in kalendarischen Einheiten verknüpft, und diese Daten müssen als reine Fakteninformation ständig geübt und eingepägt werden. Auch besteht keine gesamte chronologische Ordnung der Repräsentationen. Die zeitliche Dimension der autobiografischen Erinnerung entsteht vielmehr beim Abruf durch *rekonstruktive Inferenzprozesse* (Friedman, 1993; Larsen u.a., 1996). Diese beruhen vor allem auf zwei Mechanismen: der Auswertung von *temporalen Abruf-Hinweisreizen* und der Etablierung von *order codes*.

Temporale Abruf-Hinweisreize sind inhaltliche Bestandteile der Repräsentation oder der Verbindungen von Repräsentationen, aus denen unter Rückgriff auf allgemeines Wissen um Zeitläufe und Abfolgen inferiert werden kann, wann ein Ereignis stattgefunden haben kann und wann es nur sehr unwahrscheinlich stattgefunden haben kann (Friedman, 1993; Larsen u.a., 1996). Dies kann Wissen über das individuelle Leben und seine zeitlichen Abfolgen und Rhythmen sein – eine Begegnung mit einem bestimmten Kollegen kann dem Lebensabschnitt in einer bestimmten Firma zugeordnet werden, der Erinnernde ist jedes Jahr zu Ostern nach Italien gereist – oder allgemeines Wissen über Zeitabläufe. So kann über die Erinnerung des Wetters am Umzugstag zumindest die Jahreszeit des Umzugs erschlossen werden; man weiß, dass ein Hochschulstudium im Allgemeinen mindestens viereinhalb Jahre dauert und deshalb schwerlich bereits zwei Jahre nach Aufnahme mit dem Diplom abgeschlossen werden kann usw.

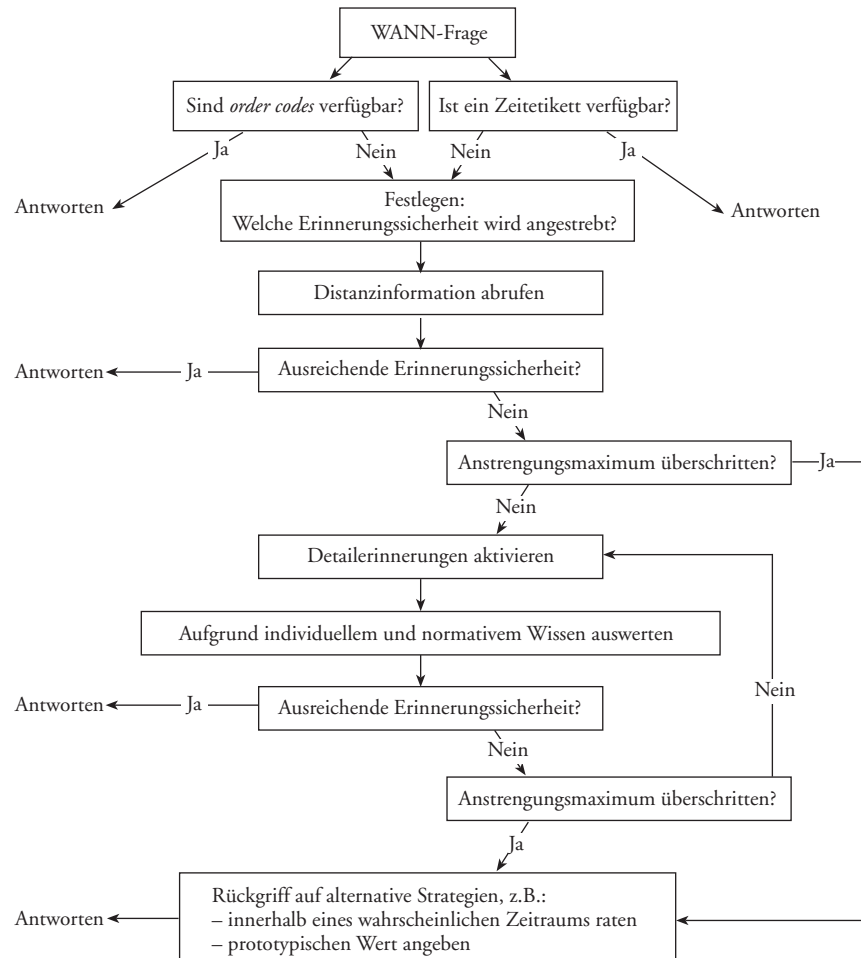
Die zweite Informationsquelle, die *order codes*, sind Verweise einer Information auf eine andere und entstehen, wenn innerhalb der narrativen und um die zentralen Selbstaspekte herum organisierten lokalen Ordnung der autobiografischen Gedächtnisrepräsentationen eine zeitliche Sequenz entsteht (Brown, Rips, & Shevell, 1985; Brown & Schopflocher, 1998; Friedman, 1993). *Order codes* bestehen nicht zwischen allen Ereignissen, sondern zwischen solchen, die in einer bedeutsamen Beziehung stehen, zum Beispiel einer kausalen oder instrumentellen Beziehung („Ich bin von A. nach B. gezogen, WEIL ich den Job bei Firma F. in B. bekam“ bedeutet „Ich bin umgezogen, ALS ich den Job bei Firma F. in B. bekam“ oder „Ich habe in A. gelebt, BIS ich den Job bei Firma F. in B. bekam“). Insofern spiegeln auch *order codes* die inhaltliche Organisation des autobiografischen Gedächtnisses wider.

Einige ausgewählte Erinnerungen als so genannte Grenzsteine (*landmark events*) haben zentrale Bedeutung für die zeitliche Strukturierung des autobiografischen Gedächtnisbestands, das heißt, *order codes* werden bevorzugt mit den Grenzsteinereignissen geschaffen (z.B. Fromholt & Larsen, 1994). Shum (1998) nennt drei Voraussetzungen dafür, dass ein Ereignis als Grenzstein bezeichnet werden kann: *persönliche Beteiligung* des Erinnernden, *persönliche Bedeutsamkeit* und eine *strukturierende Rolle* in der (subjektiven) Lebensgeschichte. Besonders solche Ereignisse werden zu Grenzsteinen, die folgenschwere Veränderungen und Übergänge in den Lebensumständen und sozialen Rollen einer Person mit sich bringen – etwa Geburt eines Kindes, Ausbildungsbeginn oder der Tod einer nahe stehenden Person, aber auch öffentliche Ereignisse wie ein Kriegs- oder Vulkanausbruch, wenn diese die Person direkt betroffen haben (Auriat, 1996; Brown, Shevell, & Rips, 1986; Friedman, 1993; Fromholt & Larsen, 1994; Larsen u.a., 1996; Lieury, Aiello, Lepreux, & Mellet, 1980; Loftus & Marburger, 1983)<sup>3</sup>.

<sup>3</sup> Eine Taxonomie relevanter Eigenschaften von Ereignissen, die den Lebensverlauf gestalten und strukturieren, findet sich bei Baltes und Baltes (1990).



Abbildung 2.3: Strategien zur zeitlichen Einordnung und Rekonstruktion eines Kalenderdatums



Quelle: Friedman, 1993.

Eine solche zeitliche Struktur von aufeinander bezogenen Repräsentationen kommt dabei noch ohne einen allgemeinverbindlichen Kalender aus. Die Vergabe eines Kalenderdatums erfolgt dann über eines der wenigen Ereignisse, dessen kalendarisches Datum gespeichert wird (z.B. Geburtstage, Weihnachten). Durch solche Referenzereignisse kann das Datum des gesuchten Ereignisses erschlossen oder wenigstens näherungsweise eingegrenzt werden („nach Weihnachten, aber vor Sylvester“). Die Verankerung mit dem Kalender erfolgt dabei getrennt auf verschiedenen Zeitskalen: Der Wochentag etwa wird unabhängig vom Monat rekonstruiert, das Jahr unabhängig vom Monat (Friedman & Wilkins, 1985; Larsen u.a., 1996).

Neben dieser relativ arbeitsaufwändigen Strategie der Datierung existieren auch noch andere, weniger arbeitsintensive: Zeitliche Einordnungen werden unter anderem vorgenommen, indem die Distanz zur Gegenwart geschätzt, der Zeitabstand aus der Erinnerungssicherheit geschätzt, innerhalb eines wahrscheinlichen Zeitraums geraten oder ein prototypischer Wert angegeben wird („am ersten Januar“ oder „nach einem halben Jahr“) (Huttenlocher, Hedges, & Bradburn, 1990). Abbildung 2.3 zeigt ein Modell von Friedman und Wilkins (1985), wie solche strategischen Schritte aufeinander aufbauen und einander ergänzen können.

Die zeitliche Einordnung von Erinnerungen und die Vergabe kalendarischer Daten sind also vom inhaltlichen Erinnern getrennt ablaufende Prozesse, die aber ebenfalls auf die schematischen Strukturen der Erinnerungen und ihrer Verbindungen untereinander beruhen. Besonders die Vergabe von Kalenderdaten ist dabei eine relativ hohe Anforderung, die den Strukturen des autobiografischen Erinnerns nicht selbstverständlich innewohnt. Deshalb kann es auch beim Datieren zu Abweichungen kommen, die auf der Grundlage der Organisationsstrukturen erklärt werden können.

## 2.5 Fehler beim autobiografischen Erinnern: Ein Überblick über die Befundlage

### 2.5.1 Vergessensanfällige Information

Welche autobiografische Information kann zeitstabiler und korrekter rekonstruiert werden, welche hat ein hohes Risiko, vergessen oder falsch rekonstruiert zu werden, und wie unterscheiden sich rekonstruierte Biografien von tatsächlichen? Zum einen sind autobiografische Fakten ohne Erlebnisgrundlage, die per definitionem keine Verknüpfungen haben können, besonders vergessens- und instabilitätsanfällig. Fakteninformation kann über längere Zeit *nur* durch immer wiederholten Wiederabruf und Neueinspeicherung im Langzeitspeicher konsolidiert werden. Wenn sie nicht replikativ erinnert werden kann, ist es nur schwer möglich, sie inferenziell zu rekonstruieren<sup>4</sup>. Für alle Zahlen, Daten und Fakten muss also ein besonders hohes Risiko des Vergessens oder Fehlerinnerns angenommen werden.

Für selbst Erlebtes ist zu erwarten, dass die Informationen besser erinnert und rekonstruiert werden können, die als Episoden zentrale Bestandteile der Repräsentationsstruktur sind oder sich aus diesen rekonstruieren lassen. Dies sind Inhalte, die für das Selbst bzw. für die narrative Struktur des Lebenslaufs bedeutsam sind. Verschiedene Ereignisseigenschaften wurden untersucht, inwieweit sie die Behaltensgüte beeinflussen, darunter *Folgenschwere* (*consequentiality*), *Einzigartigkeit* und *emotionaler Gehalt*.

Folgenschwere Ereignisse, die Konsequenzen für das Leben der Erinnernden haben, werden besser erinnert als folgenlose „Intermezzi“ (Barrick, Hall, & Berger, 1996; Brewer, 1986;

---

<sup>4</sup> Die Forschung zu Mnemotechniken befasst sich mit der Frage, durch welche bewusst eingesetzten „Tricks“ und „Eselsbrücken“ solche bedeutungsarmen Inhalte zuverlässiger gespeichert werden können. Etwa wurde das Herstellen von „Pseudobedeutungen“ beschrieben (Ericsson & Chase, 1982), das Erzeugen visueller Repräsentationen (Bower & Reitman, 1972) sowie die Methode, dass sie mit einem eigens ausgewählten Schema in Verbindung gesetzt werden (Groninger, 1971; Kliegl, Smith, Heckhausen, & Baltes, 1987). Insgesamt laufen alle diese Techniken darauf hinaus, die Inhalte mit den für rekolllektive Erinnerungen typischen Eigenschaften (Bedeutung, Bildhaftigkeit, Kontexte) anzureichern und so die schematische Speicherstruktur doch für die faktische Information nutzbar zu machen.

Rubin & Kozin, 1984), da sie eher Bestandteil der Lebensgeschichte werden (müssen) und insofern auch häufiger rekonstruiert und geübt werden. Einzigartige, unverwechselbare Ereignisse werden besser erinnert als ähnliche, die in Serie auftreten und deshalb eher als generische Episoden repräsentiert werden (Thompson u.a., 1996), innerhalb derer distinkte Details verblasen und die Verwechslungsgefahr hoch ist.

Komplex ist die Rolle der emotionalen Bedeutsamkeit für die Erinnerungsgüte. Manche Studien zeigen, dass erhöhte emotionale Intensität zu besserem Behalten führt, insbesondere die Untersuchungen zu den so genannten „Blitzlichterinnerungen“. Hier erinnern sich die Befragten auch noch nach langer Zeit detailreich und lebhaft an emotional aufgeladene Momente – etwa was sie taten, als sie von der Ermordung des US-amerikanischen Präsidenten Kennedy erfuhren (Brown & Kulik, 1977) oder von der Explosion der Raumfähre „Challenger“ (Neisser & Harsch, 1992). Möglicherweise sind diese Blitzlichterinnerungen aber nicht so korrekt, wie die Studienteilnehmer behaupten und empfinden (Neisser & Harsch, 1992). Auch vergessen die Teilnehmer in einer Untersuchung von Wagenaar und Groeneweg (1990), in der die Erinnerungen von Überlebenden von Konzentrationslagern direkt nach der Befreiung und 40 Jahre später verglichen werden, sogar schwere Misshandlungen oder das Mitansehenmüssen von Morden. Andere Studien belegen, dass nicht der Grad, sondern die Polarität der Emotion ausschlaggebend ist: Angenehme Ereignisse werden besser als unangenehme und unangenehme besser als neutrale erinnert (Linton, 1982; Wagenaar, 1986). Christianson und Loftus (1991) vermuten, dass emotional aufgeladene Ereignisse zwar seltener vergessen werden als neutrale, dass aber weniger Detailinformation korrekt erinnert werden kann. Häufiger ist dabei die Verzerrung in eine Richtung, die die Person als positiver erscheinen lässt; etwa werden vergangene Leistungen eher überhöht und vergangene Niederlagen heruntergespielt (Barrick u.a., 1996). Dies ist jedenfalls für psychisch Gesunde der Fall; depressiv Erkrankte zeigen eine umgekehrte Verzerrung im Einklang mit dem aktuellen ungünstigen Selbstbild und der schlechten Stimmung und erinnern eher negative, abwertende Ereignisse (Sheppard & Teasdale, 1996).

Folgeschwere, Einzigartigkeit und emotionale Bedeutsamkeit bzw. Polarität eines Ereignisses oder einer Erinnerung ergeben sich zum Teils auch erst in der Rückschau vor dem Hintergrund der jeweils aktuellen autobiografischen Schemata. Beispielsweise können Emotionen beim Erleben im Rahmen des elaborierenden Erinnerns von den Emotionen bei der aktuellen oder bei einer zwischenzeitlichen Rekonstruktion überlagert werden oder ein damals bedeutsam und folgeschwer erscheinendes Ereignis wird in der Rückschau als unwichtiges Intermezzo eingeschätzt und aus der Lebensgeschichte verbannt. Es werden also solche Ereignisse selektiv schlechter erinnert oder umgedeutet, die aus der aktuellen Sicht keine Folgen hatten, in einer Reihe ähnlicher Ereignisse auftraten und emotional negativ waren. Die rekonstruierten Erinnerungen erscheinen dadurch folgeschwerer, einzigartiger und emotional positiver, als die tatsächlichen Erlebnisse es waren.

### 2.5.2 Unterschiede zwischen Personengruppen

Im Folgenden werde ich einen Überblick über die Befundlage zur unterschiedlichen Erinnerungsgüte von Männern und Frauen sowie von unterschiedlichen Altersgruppen geben. Interessant wäre auch eine Betrachtung unterschiedlicher Bildungs- oder sozialer Statusgruppen;

hierzu liegen aber meines Wissens keine gedächtnispsychologischen Befunde vor. Für Gruppenunterschiede gibt es kognitive Erklärungen aus drei Bereichen:

- (a) die Personengruppen haben *unterschiedliche Ereignisse zu erinnern*,
- (b) die Personengruppen unterscheiden sich in ihren *kognitiven Kapazitäten*,
- (c) die Personengruppen unterscheiden sich in den die Rekonstruktion und Repräsentation anleitenden und organisierenden *Schemata*.

#### *Alter der Erinnernden*

Das autobiografische Gedächtnis entwickelt sich über den Lebenslauf durch die Akkumulation von Inhalten und ihre allmählich fortschreitende schematisch integrierte Repräsentation (Fitzgerald & Lawrence, 1984; Fivush, Haden, & Adam, 1995; Rubin, 1982; Rubin, Wetzler, & Nebes, 1986; Wetzler & Sweeney, 1986). Erst ab einem Alter von etwa 16 Monaten können Kinder Erinnerungen mit Unterstützung von Erwachsenen rekonstruieren. Aktives oder spontanes Erinnern kommt ab etwa zweieinhalb Jahren vor (Fivush, Gray, & Fromhoff, 1987). In dieser Zeit verfügen Kinder nur über ereignisspezifisches Wissen im Sinne von Conway (siehe Abschnitt 2.3.1), nicht aber über eine übergeordnete, schematische Gedächtnisorganisation. Deshalb können sich Menschen zumeist gar nicht oder nur diffus und unspezifisch an Ereignisse aus den ersten Lebensjahren erinnern, ein Phänomen, das „frühkindliche Amnesie“ genannt wird (Fitzgerald & Lawrence 1984; Fivush u.a., 1995; Rubin, 1982; Rubin u.a., 1986; Wetzler & Sweeney, 1986). Schemageleiteter Abruf ist nicht vor dem Schulalter zu erwarten; das Entstehen eines *life story*-Schemas wird sogar erst in der Adoleszenz beobachtet (Bluck & Habermas, 2000; Habermas & Bluck, 2000). Dies wird zum einen darauf zurückgeführt, dass die Entwicklung des kognitiven Apparats komplexere Integrationsleistungen erst ab einem gewissen Alter ermöglicht. Zum anderen spielen auch motivationale Gründe eine Rolle, da erst in der Adoleszenz das Entstehen einer Lebensgeschichte zur Identitätsbildung subjektiv notwendig wird.

Im höheren Alter ist die Erinnerungsgüte durch mehrere Faktoren gefährdet. Durch die Länge des Lebens müssen mehr und weiter zurückliegende Ereignisse erinnert werden. Weiterhin lässt die Leistungsfähigkeit der Informationsverarbeitung nach, vor allem im Bereich der „fluiden Intelligenz“ (Verarbeitungsgeschwindigkeit, Flexibilität, Aufmerksamkeitsbreite), weniger im Bereich der „kristallinen Intelligenz“ (organisiertes gut geübtes Wissen) (Baltes, Lindenberger, & Staudinger, 1998; Cattell, 1971). Ältere Menschen konzentrieren sich auf ausgewählte, wichtige Aspekte und konsolidieren diese durch wiederholtes Üben. Dadurch verfügen sie über einen Satz ausgewählter Erinnerungen an persönlich relevante Ereignisse, die mit hoher Detailgenauigkeit und Zuverlässigkeit erinnert werden können; weniger wichtige Erinnerungen fallen quasi einer adaptiven, durch die Ressourcenknappheit provozierten Selektionsmaßnahme zum Opfer (Baltes & Baltes, 1990; Cohen, 1998; Cohen, Conway, & Maylor, 1994; Fitzgerald & Lawrence, 1984; Rabbitt & Winthorpe, 1988). Dadurch werden vor allem Inhalte, die nicht direkt repräsentiert sind, sondern indirekt rekonstruiert werden müssten, schwieriger und schlechter erinnert. Dies betrifft vor allem das Datieren, die Erinnerung an Details und an Kontexte (Bluck, Levine, & Laulhere, 1999; Burke & Light, 1981; Cohen, 1998).

Darüber hinaus spielt mit zunehmendem Alter das Nachdenken über das eigene Leben eine zentralere und qualitativ andere Rolle (Coleman, 1991). Die Entwicklung einer sinnhaften Lebensgeschichte wird besonders wichtig für Zufriedenheit und psychische Gesundheit.

Insofern kann es bei älteren eher als bei jüngeren Menschen dazu kommen, dass sie Ereignisse, die nicht diesem überlernten Narrativ angehören, zunehmend seltener abrufen und dadurch auch zunehmend schwerer rekonstruieren können. Wenn hingegen Ereignisse abgefragt werden, die Konstituenten der Lebensgeschichte sind, kann es bei älteren Menschen eher als bei jüngeren geschehen, dass diese verzerrt rekonstruiert werden, um sich besser in das Schema der individuellen Lebenslogik zu fügen.

Es bestehen also für Ältere dieselben Rekonstruktionsschwierigkeiten wie für Jüngere, aber in stärkerem Umfang. Für die zentralen Erinnerungen der Lebensgeschichte haben die Altersunterschiede weniger Bedeutung; beim Rekonstruieren wenig geübter, wenig zentraler und zugänglicher Information hingegen sind Ältere besonders benachteiligt. Allerdings müssen Alters- und besser noch Leistungsgruppen differenziert werden. Man nimmt an, dass bis zum Alter von etwa 80 Jahren ein großer Teil der tendenziell zunehmenden kognitiven Beeinträchtigungen kompensiert werden kann. Außerdem lässt sich durch geeignete Erinnerungshilfen auch bei älteren Befragten die Erinnerungsgüte deutlich steigern. Allerdings gibt es innerhalb der Gruppe der Älteren, ebenso wie bei jüngeren Altersgruppen, deutliche Unterschiede, die nicht mit dem chronologischen Alter zusammenhängen, sondern die ursprünglichen kognitiven Unterschiede zwischen Individuen widerspiegeln (Jobe, Keller, & Smith, 1996; Mayer u.a., 1996; Reischies & Lindenberger, 1996). Auch die zunehmende Verbreitung von klinischen Demenzsymptomen erschwert in der Gruppe der „alten Alten“ das zuverlässige Erinnern und Datieren (Fromholt & Larsen, 1994; Helmchen u.a., 1996; Maas, Borchelt, & Mayer, 1996).

### *Geschlecht*

In vielen Untersuchungen erweisen sich Frauen sowohl beim Erinnern als auch beim Datieren von Ereignissen als überlegen, die frühkindliche Amnesie ist bei ihnen schwächer ausgeprägt, und ihre Erinnerungen sind lebhafter, elaborierter und emotionaler als die von Männern (Friedman & Pines, 1991; Ross & Holmberg, 1990; Skowronski & Thompson, 1990). Dies wird auf die unterschiedliche Sozialisation und das unterschiedliche Kommunikationsverhalten der Geschlechter zurückgeführt: Frauen erzählen ihre Erinnerungen häufiger und üben und strukturieren sie dadurch besser. Geschlechterunterschiede müssen allerdings differenziert nach Art der Erinnerungen und nach Art der Erinnerungsfehler betrachtet werden. Beispielsweise belegen Thompson u.a. (1998), dass Frauen beim Datieren zwar seltener ein abweichendes Datum produzieren, dass ihre Abweichungen aber genauso groß und in dieselbe Richtung ausfallen wie die von Männern. Und für die Erinnerung aus dem Bereich Arbeit und Erwerbstätigkeit, wie sie vor allem in der Befragungsmethodologie betrachtet wurden, machen Frauen und ganz besonders junge Frauen fehlerhaftere Angaben als Männer, sowohl zu Phasen der Erwerbstätigkeit als auch der Arbeitslosigkeit (Dex & McCulloch, 1997; Elias, 1997; de Graaf & Wegener, 1989; Thélot, 1990). Dies wird darauf zurückgeführt, dass Frauen, besonders jüngere Frauen, einerseits wechselhaftere Erwerbskarrieren mit vielen Teilzeitbeschäftigungen und Phasen der Heim- und Erziehungsarbeit zu rekonstruieren haben und dass andererseits die Erwerbssphäre in ihrem Selbst und ihrem Lebensnarrativ eine geringere Rolle spielt und deshalb weniger organisiert repräsentiert wird. Ein interessanter Befund von Brückner (1993) zeigt, dass Frauen, die Kinder hatten, ihre Berufslaufbahn besser erinnern als kinderlose Frauen, während die Erinnerung männlicher Teilnehmer nicht davon beeinflusst wird, ob sie Kinder hatten oder nicht. Dies kann so interpretiert werden, dass die

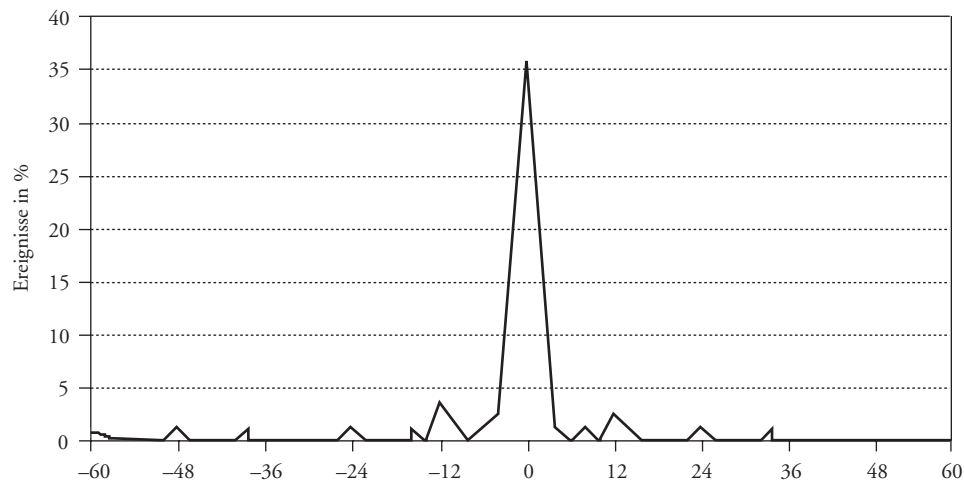
Geburten und Geburtstage von Kindern für Frauen eine sehr klare inhaltliche und zeitliche Organisationsstruktur und Orientierungshilfe beim Erinnern bieten.

### 2.5.3 Fehler beim Datieren

Wie in Abschnitt 2.4 dargelegt, ist die raumzeitliche Lokalisiertheit ein zentrales Charakteristikum autobiografischer (episodischer) Information, und die zeitliche Einordnung von Erinnerungen in das Bezugsgefüge des autobiografischen Gedächtnisses erfolgt auf der Grundlage ganz ähnlicher Mechanismen wie das Erinnern selbst. Die Vergabe von Kalenderdaten hingegen ist eine Leistung, die nicht in der Struktur des autobiografischen Erinnerns eingebettet ist. Deshalb bereitet auch das ungefähre zeitliche Einordnen von Erinnerungen – etwa bestimmter Episoden innerhalb bestimmter Lebensabschnitte oder eine Episode zeitlich vor einer anderen – im Allgemeinen keine größeren Probleme, beim kalendarischen Datieren kommt es aber sehr häufig zu Gedächtnisfehlern (Friedman, 1993).

Am korrektesten und zeitstabilsten werden natürlich die Ereignisse erinnert, die mit einem Zeitetikett zusammen repräsentiert sind: Wenn die Teilnehmer „einfach wussten“, zu welchem Datum ein Ereignis stattfand, sind die Abweichungen selten und sehr gering (Burt, 1992; Lieury u.a., 1980; Thompson, Skowronski, & Lee, 1987, 1988). Dies trifft aber nur auf wenige Ereignisse zu, vor allem für Ereignisse wie Geburtstage oder Eheschließungen und für sehr kurz zurückliegende Ereignisse. Für die Mehrzahl aller Erinnerungen muss das Datum inferenziell rekonstruiert werden.

Abbildung 2.4: Skaleneffekt bei der Rekonstruktion eines Kalenderdatums



Abweichung der Befragtenangaben zum Datum des ersten Umzugs von den Einträgen im belgischen Melderegister: symmetrische Verteilung und Häufung im Jahresabstand (Skaleneffekt).

Quelle: Auriat, 1996.

Für die Rekonstruktion von Daten gilt dasselbe wie für die Rekonstruktion allgemein: Folgeschwere, emotional bedeutsame und einzigartige Ereignisse, die fest in die temporale Struktur des autobiografischen Gedächtnisses eingebunden und detailreich und kontextualisiert repräsentiert sind, werden besser und zeitstabiler datiert, da sie mehr temporale Abruf-Hinweisreize und *order codes* aufweisen (Thompson u.a., 1996).

Wenn Fehler beim Datieren gemacht werden, können diese größer oder kleiner ausfallen und sich entweder mehr oder weniger gleichmäßig um das korrekte Datum verteilen. Zahlreiche Befunde sprechen dafür, dass die meisten Datierungen nur einige wenige Zeiteinheiten daneben fallen und dass abweichende Daten gleich häufig zu früh und zu spät angegeben werden (Larsen & Thompson, 1995; Thompson u.a., 1996). Im Folgenden möchte ich zwei Fehldatierungsformen besprechen, die sich als Ergebnis der Rekonstruktionsprozesse erklären lassen: Skaleneffekte und Teleskopieren. Besonders zum Teleskopieren gibt es umfangreiche Studien gerade aus Befragungskontexten, weshalb ich dieses Phänomen und die Bedingungen seines Auftretens ausführlich und kritisch diskutieren werde.

### *Skaleneffekte*

In einigen Studien wurde nachgewiesen, dass sich unzutreffende Zeitangaben nicht zufällig über den zur Verfügung stehenden Zeitraum verteilen. In der „Enquête 3B.bis“ werden die Daten der Umzüge der Befragten mit den Daten des belgischen Melderegisters verglichen. In Abbildung 2.4 sehen wir, dass sich die Abweichungen der Befragtenangaben von den Registerinträgen bei 10 bis 13 Monaten, 22 bis 26 Monaten und 34 bis 37 Monaten vor und nach dem wahren Datum häufen (Auriat, 1993, 1996) – also etwa im Jahresabstand. Die Datierung auf der Monatsskala ist also recht genau, nicht aber die auf der Jahresskala<sup>5</sup>. Ähnliche Häufungen berichten Rubin und Baddeley (1989) sowie Friedman und Wilkins (1985). Thompson (1998) belegt in einer Reihe von Tagebuchstudien Häufungen der Abweichungen im Wochenabstand, das heißt, auf der Wochentagsskala wird die zeitliche Einordnung korrekt vorgenommen, nicht aber auf der Wochenskala.

Dieses Verteilungsphänomen nennen Friedman und Wilkins „Skaleneffekt“, denn es entsteht dadurch, dass die Rekonstruktion eines Datums auf verschiedenen Skalen des Kalenders unabhängig voneinander und aufgrund anderer temporaler Hinweisreize und *order codes* vorgenommen wird (siehe Abschnitt 2.4). So kann der Monat eines Umzugs etwa gut aus den Wetterbedingungen erschlossen werden, die aber keinen Rückschluss auf das Jahr zulassen; dieses wird aus dem Sterbedatum der im selben Jahr gestorbenen Großmutter abgeleitet. Wenn sich ein Irrtum nur auf der Jahresskala einschleicht, weicht das erinnerte Umzugsdatum um genau ein oder mehrere Jahre vom tatsächlichen Datum ab. Ein solcher Irrtum ist offenbar denjenigen Teilnehmern unterlaufen, die die Fehlerhäufungen in Abbildung 2.4 verursacht haben. Ob die Datierung einer Erinnerung auf einer Skala besser gelingt als auf einer

<sup>5</sup> Die sehr häufigen Abweichungen innerhalb der ersten drei Monate lassen sich so allerdings nicht erklären. Diese wurden von der Autorin darauf zurückgeführt, dass das belgische Gesetz eine Frist zur Ummeldung von bis zu acht Tagen nach dem Umzug vorsieht und dass von der Abgabe des entsprechenden Formulars bis zur Änderung des Eintrags im Register von einer maximal dreimonatigen Bearbeitungsfrist ausgegangen werden kann.

anderen, und welche dies ist, hängt auch von dem Ereignis selbst ab (Larsen & Thompson, 1995). Für Umzüge sind Monate offenbar leichter zu rekonstruieren als Jahresangaben.

### *Teleskopieren*

In der Literatur findet sich häufig der Hinweis auf eine Neigung des autobiografischen Erinnerns, abweichende Datierungen tendenziell so vorzunehmen, dass ein weniger lang zurückliegendes Datum genannt wird. Dies wird als „Teleskopieren“ bezeichnet (Neter & Wakesberg, 1964), da Ereignisse quasi in der Rückschau wie mit einem Teleskop näher an die Gegenwart herangeholt werden.

Tatsächlich handelt es sich beim Teleskopieren aber um einen Fehler bei der Häufigkeitsschätzung, wenn Personen angeben sollten, wie oft sie ein bestimmtes Ereignis in einer bestimmten Referenzperiode erlebt hatten – etwa, wie oft sie in den vergangenen sechs Monaten beim Arzt gewesen oder wie häufig sie in den letzten zehn Jahren umgezogen seien. Explizit oder implizit nehmen viele Autoren an, dass die Teilnehmer solche Fragen beantworten, indem sie sich zuerst an alle relevanten Ereignisse und Aktivitäten *erinnern*, diese dann (mit dem zu erwartenden Fehler) *datieren und zuletzt zählen*, wie viele davon in die relevante Periode fallen (Loftus & Marburger, 1983; Rubin & Baddeley, 1989; Thompson u.a., 1988). Innerhalb dieses Datengenerierungsmodells können fehlerhafte Häufigkeitsangaben als Resultat von Fehldatierungen erklärt werden: Wenn die Häufigkeit überschätzt wurde (anstelle, wie zu erwarten gewesen wäre, durch Vergessen unterschätzt), wurde dies so interpretiert, dass diese Teilnehmer irrtümlich auch Ereignisse aus der Zeit vor der abgefragten Periode mit erinnert, in den Referenzzeitraum hinein „teleskopiert“ und beim Zählen berücksichtigt hatten (Bradburn, Rips, & Shevell, 1987; Loftus & Marburger, 1983; Neter & Wakesberg, 1964; Silberstein 1989; Sudman & Bradburn, 1973; Thompson u.a., 1988)<sup>6</sup>. Teilweise wird auch das Unterschätzen von Häufigkeiten mit einer umgekehrten Neigung „Rückwärtsteleskopieren“ erklärt – die Teilnehmer hätten hier irrtümlich Ereignisse aus der abgefragten Periode aus dieser hinaus verlegt (Sudman & Bradburn, 1973).

Loftus und Marburger (1983) nehmen eine verzerrende Eigenschaft der menschlichen Wahrnehmung als Ursache des Teleskopierens an: die „Zeitkompression“, eine Neigung zur Unterschätzung der seit einem Ereignis verstrichenen Zeit. Das Rückwärtsteleskopieren wird

<sup>6</sup> Dreierlei Studien verstehen sich als Belege für einen Teleskopeffekt:

- Die Häufigkeitsangaben bestimmter Ereignisse für unterschiedlich lange Zeiträume werden verglichen (etwa die Angaben für die letzten sechs und die letzten zwölf Monate). Dabei berichten Studienteilnehmer mit längeren retrospektiven Perioden mehr Ereignisse pro Zeiteinheit; dies wird als Hinweis darauf interpretiert, dass sie irrtümlich auch Ereignisse aus der Zeit vor der abgefragten Periode mit berücksichtigt haben (z.B. Bradburn u.a., 1987; Sudman & Bradburn, 1973; Thompson u.a., 1988).
- Die Häufigkeitsangaben bestimmter Ereignisse von Personen, die mit herkömmlichen Methoden befragt wurden, werden mit solchen verglichen, bei denen mit elaborierten Befragungstechniken ein Fehldatieren verhindert werden sollte. Da Letztere wirklich weniger Ereignisse angeben, nehmen die Autoren an, dass Erstere irrtümlich auch Ereignisse aus der Zeit vor der abgefragten Periode mit berücksichtigt haben, was bei Letzteren wie geplant verhindert worden war (z.B. Loftus & Marburger, 1983; Neter & Wakesberg, 1964).
- Silberstein (1989) bat seine Teilnehmer, die exakten Daten der von ihnen berücksichtigten Ereignisse innerhalb der abgefragten Periode anzugeben. Davon fallen disproportional viele in den Zeitraum näher zum Interview und disproportional weniger in den länger zurückliegenden Zeitraum. Auch dieser Umstand wird so gedeutet, dass ein Teil der Ereignisse irrtümlich vordatiert wurde.



von Sudman und Bradburn (1973) auf das Zusammenspiel von Vergessen und Zeitkompression zurückgeführt: die Neigung zur Zeitkompression fördert das Zuviel-Berichten von Ereignissen, die Neigung zum zunehmenden Vergessen im Verlauf der Zeit fördert das Zuwenig-Berichten, und je nach der relativen Größe der beiden Neigungen wird die Häufigkeit insgesamt überschätzt (wenn die Zeitkompression das Vergessen überwiegt), unterschätzt (wenn das Vergessen die Zeitkompression überwiegt) oder korrekt angegeben (wenn Zeitkompression und Vergessen sich gegeneinander aufheben).

Rubin und Baddeley (1989) hingegen bestreiten, dass Über- oder Unterschätzungen von Häufigkeiten als Beleg für systematische Fehler bei der Datierung und einer verzerrten Zeitrepräsentation gelten können. Erinnernde wenden auch andere Strategien bei der Beantwortung von Häufigkeitsfragen an, als tatsächlich alle Ereignisse zu erinnern, zu datieren und die relevanten zu zählen: Sie stellen grobe Schätzungen an, nennen einen für „durchschnittlich“ gehaltenen Wert oder vergleichen die Vergangenheit mit der Gegenwart als Bezugsgröße (Huttenlocher u.a., 1988; Sudman, Bradburn, & Schwarz, 1996; Thompson u.a., 1988)<sup>7</sup>. Auch die Neigung der Befragten, im Zweifelsfall lieber ein Ereignis zuviel zu berichten als eines zu wenig, wird als eine (motivationale) Erklärung angeboten (Rubin & Baddeley, 1989).

Viele Studien, in denen die Datierung einzelner Ereignisse mit Registern oder Tagebucheinträgen verglichen werden, finden überhaupt keine generelle systematische Neigung in irgendeine Richtung (Auriat, 1993; Rubin & Baddeley, 1989; Thompson u.a., 1998; Wagenaar, 1986; White, 1982). Andere belegen hingegen eine systematische Neigung, ein späteres, näher an die Gegenwart herangerücktes Datum anzugeben (Huttenlocher u.a., 1988; Johnson, Gerstein, & Rasinski, 1998; McGovern, Lurie, Margolis, & Slater, 1998; Thompson u.a., 1988). Thompson u.a. (1988) nehmen an, dass die Länge der retrospektiven Distanz eine Rolle dabei spielt, ob Teleskopieren auftritt und in welche Richtung: Sie belegen systematische Fehldatierungen erst für Zeiträume, die länger als etwa acht Wochen waren. Bei retrospektiven Perioden von einigen Wochen und Monaten kommt es dabei eher zu vorwärts gerichtetem, für mehrjährige Intervalle zu rückwärts gerichtetem Teleskopieren. Wagenaar (1986) berichtet einen Einfluss der emotionalen Intensität: In seiner introspektiven Tagebuchstudie hatte er Ereignisse, die er als „stark emotional“ einstufte, im Allgemeinen näher an die Gegenwart herandatiert, das heißt die seither verstrichene Zeit unterschätzt. Thompson u.a. (1996) stellen hingegen fest, dass Ereignisse nicht alle in dieselbe Richtung fehldatiert, sondern tendenziell weg von den Rändern der Referenzperiode und hin zu deren Mitte verlegt wurden – dass also *länger zurückliegende* Ereignisse eher weiter nach vorne an die Gegenwart herangerückt wurden und dass *kürzer zurückliegende* Ereignisse eher weiter in die Vergangenheit hineinverlegt wurden; Anhaltspunkte hierfür finden sich auch bei Brown u.a. (1985) und White (1982). Dies erklären Rubin und Baddeley (1989) dadurch, dass falsche Datumsangaben zwar prinzipiell symmetrisch um die wahren Daten streuen, dass aber die Grenzen der jeweiligen Referenzperiode auch Grenzen für die Fehldatierung setzen: Ereignisse, die sehr weit zurückliegen, können, wenn sie fehldatiert werden, nur näher zur Gegenwart verschoben werden, da sie sonst aus der Referenzperiode herausfallen würden. Dies fällt dem Befragten dann auf und er korrigiert seine Datumsangabe. Umgekehrt können die jüngsten Ereignisse nur zurückdatiert werden, da sie sonst in die Zukunft hineindatiert

<sup>7</sup> Welche Abrufstrategie die erinnernden Personen bei der retrospektiven Häufigkeitsschätzung anwenden, hängt unter anderem damit zusammen, wie lang die Periode ist, wie viele Ereignisse eine Person erinnern muss und wie viel Wert sie auf höchstmögliche Genauigkeit legt (Schwarz & Sudman, 1996).

werden müssten (jedenfalls in Befragungen, deren Referenzperiode bis zur Gegenwart reicht), was dem Befragten wiederum auffallen und zu einer Korrektur führen würde.

Schließlich werden auch nichtkognitive Erklärungen für systematische Abweichungen in die eine oder andere Richtung angeboten: Auriat (1996) stellt beim Vergleich von Heiratsdaten mit standesamtlichen Registern eine leichte Tendenz fest, das Datum in den seltenen Fällen, in denen es fehldatiert wurde, näher an die Gegenwart heranzuverlegen. Dies führt sie auf eine Verwechslung der standesamtlichen Trauung mit der kirchlichen Trauung zurück, die ja in der Regel später stattfindet. Die Neigung, die Eheschließung näher an die Gegenwart heranzurücken, ist also nicht durch das Datieren, sondern durch die Erinnerung des falschen Ereignisses zu Stande gekommen.

Insgesamt sind Abweichungen bei der Vergabe eines kalendarischen Datums also verbreitet, in der Regel aber zumindest für zentrale, folgenschwere und emotional bedeutsame Ereignisse gering. Abweichungen treten unabhängig voneinander auf verschiedenen Zeitskalen auf. Auf welcher Skala Fehler wahrscheinlicher sind, hängt davon ab, welche Zeiteinheiten für das jeweilige Ereignis aus temporalen Hinweisreizen oder *order codes* besser rekonstruiert werden können. Wenn Kalenderdaten fehlerhaft rekonstruiert werden, weichen sie symmetrisch in beide Richtungen vom tatsächlichen Datum ab. Durch die Begrenzung der Referenzperiode oder dadurch, dass tatsächlich ein anderes Ereignis datiert wird, treten auch systematische Abweichungen auf, und Ereignisse werden häufiger näher an die Gegenwart, weiter in die Vergangenheit oder hin zur Mitte der Referenzperiode datiert.

Die in den vergangenen Abschnitten beschriebenen Fehler beim Datieren und Rekonstruieren von Ereignissen können für jede Erinnerung auftreten, unabhängig ob sich ein Ereignis zwei oder zehn Jahre in der Vergangenheit zugetragen hat. Im folgenden Abschnitt werde ich betrachten, ob eine zunehmende Länge der Rückschau die Güte und Zeitstabilität der autobiografischen Rekonstruktion auch zunehmend verschlechtert.

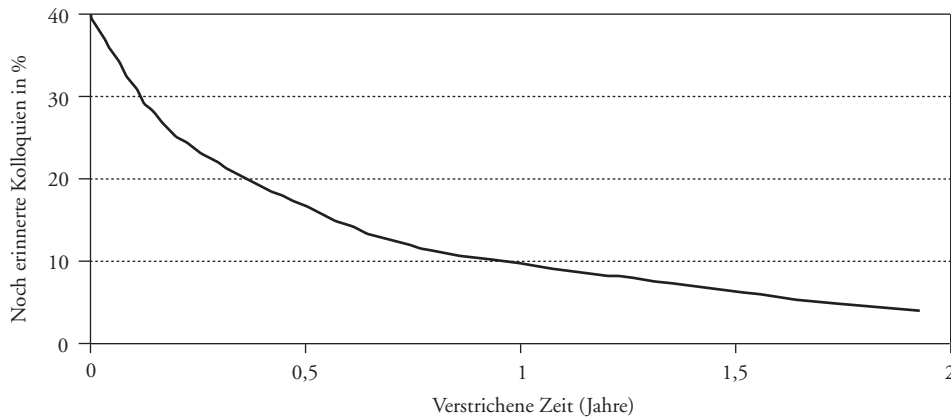
## 2.6 Vergessen, Fehldatieren und die retrospektive Distanz

### 2.6.1 Vergessen und retrospektive Distanz

Aus mehreren Gründen ist anzunehmen, dass länger zurückliegende Ereignisse weniger korrekt erinnert werden können. Thompson u.a. (1996) nehmen an, dass Erinnerungen mit zunehmender retrospektiver Distanz zum Erlebten immer seltener direkt replikativ erinnert werden können und umso häufiger rekonstruiert werden müssen. Dadurch kann es eher zu Abweichungen von der ursprünglichen Erinnerung kommen. Weiterhin werden mit der Zeit Ereignisdetails ebenfalls vergessen, wodurch unterstützende Abruf-Hinweisreize verloren gehen. Weiterhin steigt bei längeren Zeiträumen zwischen Erleben und Erinnern das Risiko, dass neue Inhalte die ursprüngliche Repräsentation im Rahmen des elaborierenden Erinnerns verändern.

Da die Rekonstruktion durch die aktuellen Selbstschemata und die aktuellen subjektiven Entwicklungslogiken angeleitet wird, können Informationen nicht mehr rekonstruiert werden, wenn sich die Ordnungsstrukturen geändert haben. Zwar sind die Schemata relativ stabil und änderungsresistent, nicht aber unveränderlich; und in verschiedenen Rekonstruktionssituationen können unterschiedliche Aspekte oder Subschemata aktiv sein (Robinson &

Abbildung 2.5: Zunahme des Vergessens mit längerer retrospektiver Distanz



Über zwei Jahre hinweg nimmt der Anteil noch erinnerter Fachbereichskolloquien kontinuierlich ab; die Abnahme ist eine kurvilineare Funktion der retrospektiven Distanz.

Quelle: Rubin & Baddeley, 1989.

Taylor, 1998). Mit längeren retrospektiven Distanzen wird es wahrscheinlicher, dass sich die Ordnungsprinzipien ändern oder dass beim Rekonstruieren unterschiedliche Selbstaspekte im Vordergrund stehen.

Wie genau sieht über die Zeit hinweg die Beziehung zwischen Vergessen und zunehmender retrospektiver Distanz aus, welche Form nimmt die „Vergessenskurve“ an? Rubin und Baddeley (1989) prüfen, wie gut sich die Mitglieder eines universitären Fachbereichs an die Fachbereichskolloquien der letzten zwei Jahre erinnern. Mit zunehmendem retrospektivem Intervall sinkt der Anteil erinnerter Kolloquien *kontinuierlich* und *kurvilinear* ab, zunächst stärker, dann schwächer (siehe Abb. 2.5).

Rubin und Wenzel (1996) ermitteln bei einer Reanalyse von mehr als 15 Studien zur autobiografischen Erinnerungsgüte für retrospektive Perioden bis zu 18 Jahren eine Gütefunktion ( $y = ax^b$ ) als die Kurve mit dem besten Fit. Dies bedeutet, dass die Erinnerungsgüte in der Zeit direkt nach dem Erleben kaum, dann aber rapide nachlässt. Ab einem gewissen Zeitpunkt reduziert sich die Geschwindigkeit der Verschlechterung wieder deutlich. Dex und McCulloch (1997) sowie Elias (1997) belegen für Phasen der Arbeitslosigkeit ein Nachlassen der Erinnerungsgüte erst ab etwa drei Jahren. Solche Verläufe könnten den Übergang von replikativem zu rekonstruktivem Erinnern widerspiegeln oder eine Veränderung der Rekonstruktionsschemata. Bei einer Untersuchung über das Erinnern an Namen und Gesichter kommen Bahrick, Bahrick und Wittlinger (1975) zu dem Schluss, dass zwar die ersten drei bis fünf Jahre nach Einspeicherung mehr und mehr vergessen wird, dass aber die Inhalte, die danach noch erinnert werden, bis zu 50 Jahre intakt bleiben<sup>8</sup>.

<sup>8</sup> In dieser Untersuchung mussten die Teilnehmer die fraglichen Inhalte allerdings wiedererkennen, anstatt sie frei zu produzieren.

Hier muss besonders zwischen *Korrektheit* und *Zeitstabilität* der Rekonstruktionen unterschieden werden. Wenn abweichende Rekonstruktionen durch die zeitliche Distanz zwischen Erleben und Rekonstruieren bedingt werden und die Rekonstruktionen dabei zunehmend schlechter ausfallen, unterscheiden sich spätere Rekonstruktionen auf charakteristische Weise von früheren Rekonstruktionen, die dem erlebten Sachverhalt noch ähnlicher sind. Durch die Tendenz der schematischen Speicherstruktur, sich zu konsolidieren, können Rekonstruktionen aber wieder und wieder übereinstimmend ausfallen, ohne jemals korrekt zu sein, wenn bereits bei den früheren Rekonstruktionen selektiv unpassende Details vergessen oder umgedeutet wurden. Dies betrifft vermutlich sogar in besonderem Maße die zentralen, selbstrelevanten Bestandteile der Repräsentationsstruktur, bei denen am ehesten die „Notwendigkeit“ zur Umdeutung besteht und die zugleich am stärksten konsolidiert werden (Cohen u.a., 1994; Neisser & Harsch, 1992; Rabbitt & Winthorpe, 1988).

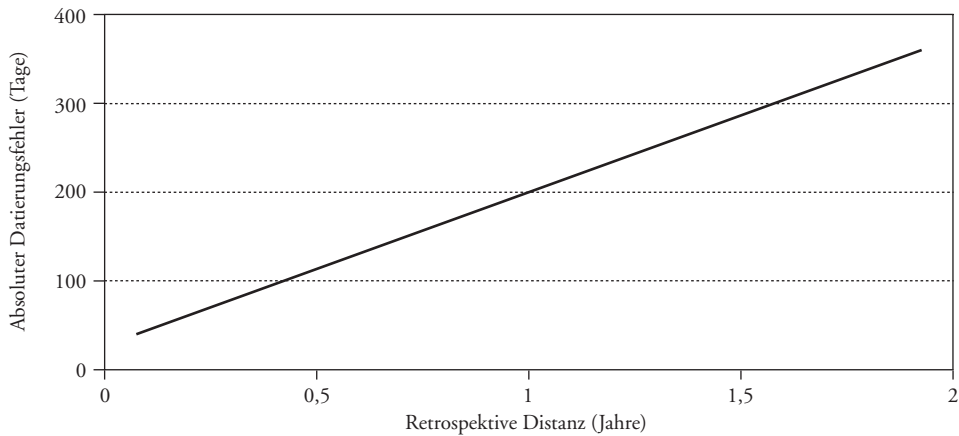
### 2.6.2 Fehldatieren und retrospektive Distanz

Auch ein Nachlassen der Datierungsgüte mit zunehmender retrospektiver Distanz lässt sich aus mehreren Gründen annehmen. Da genaue Kalenderdaten nur für eine kleine Anzahl von Ereignissen durch Wiederholung dauerhaft auswendig gewußt werden, muss für die meisten Ereignisse das Datum inferenziell rekonstruiert werden (Thompson u.a., 1996). Für länger zurückliegende Zeiträume stehen Zeitetiketten vermutlich in noch geringerem Umfang zur Verfügung. Durch das zunehmende Vergessen von Ereignisdetails sind für länger zurückliegende Erinnerungen außerdem weniger temporale Abruf-Hinweisreize verfügbar. Auch die Schemata zu deren Auswertung können sich mit der Zeit verändern, etwa wenn nach dem Übergang vom Arbeits- ins Rentenalter die Wochenstruktur (Arbeitswoche vs. Wochenende) sich verändert (Larsen & Thompson, 1995). Wiederum dürfte dies am wenigsten die festen Bestandteile der Repräsentationsstruktur betreffen, die detailreich und kontextualisiert gespeichert und über *order codes* eng ins Erinnerungsnetzwerk eingebunden sind, da hier die Abruf-Hinweisreize vergleichsweise lange erhalten bleiben.

Empirisch wurde weitestgehend übereinstimmend gezeigt, dass die Datierungsgüte tatsächlich abnimmt (Burt, 1992; Burt, Kemp, & Conway, 2001; Friedman, 1993; Linton, 1975; Rubin & Baddeley, 1989; Thompson, 1982; Thompson u.a., 1988). Allerdings hängt auch dies wieder von der Art der Inhalte ab – Umzüge, Geburtstage und Heiratsdaten etwa werden mit längerer retrospektiver Distanz nicht schlechter datiert (Auriat, 1993).

Auch die Größe der Abweichungen hängt von der Art der zu datierenden Erinnerungen ab – Fachbereichskolloquien werden zum Beispiel schon nach einem halben Jahr im Schnitt um 100 Tage abweichend datiert (Rubin & Baddeley, 1989), während die eigene Eheschließung oder Geburt eines Kindes auch Jahre und Jahrzehnte später meist nur um eine Einheit abweichend datiert wird (Auriat, 1993). Rubin und Baddeley (1989) nehmen eine lineare Zunahme der Größe der Fehldatierungen an, da die menschliche Informationsverarbeitung beim Einschätzen von zeitlichen Distanzen den Weber'schen Wahrnehmungsgesetzen gehorcht. Weber untersuchte, wie Menschen die Ausprägungen physikalischer Eigenschaften eines Objekts wahrnehmen und einschätzen (Gewicht, Lautstärke, Entfernung usw.). Er konnte belegen, dass die Schätzurteile auf regelhafte, vorhersehbare Weise von der tatsächlichen Ausprägung der jeweiligen Eigenschaft bei dem einzuschätzenden Objekt abweichen. Das Ausmaß

Abbildung 2.6: Zunahme des Datierungsfehlers mit der retrospektiven Distanz



Über zwei Jahre hinweg nehmen die numerischen Abweichungen der erinnerten von den tatsächlichen Daten der Fachbereichskolloquien kontinuierlich zu; die Zunahme ist eine lineare Funktion der retrospektiven Distanz.

Quelle: Rubin & Baddeley, 1989.

eines fehlerhaften Schätzurteils ist dabei eine Funktion der Ausprägung der einzuschätzenden Eigenschaft: Je schwerer, größer oder heller die einzuschätzenden Objekte sind, desto größer fällt auch die Abweichung der Schätzung von der wirklichen Ausprägung aus. Ein 50-Gramm-Gewicht wird vielleicht zehn Gramm zu leicht oder zu schwer eingeschätzt, ein zwei Kilogramm schweres Gewicht hingegen auch schon einmal um 100 Gramm zu leicht oder zu schwer. Die Zunahme der Abweichungen steht dabei in einer mathematisch beschreibbaren Beziehung zur Zunahme der Merkmalsausprägung. Für die physikalische Eigenschaft Gewicht etwa ist dies eine logarithmische Beziehung – wenn die Merkmalsausprägung um eine Einheit steigt, steigt der Schätzfehler um den Logarithmus dieser Einheit.

Analog hierzu meinen Rubin und Baddeley (1989), dass der Datierungsfehler eine Funktion der Größe der einzuschätzenden Eigenschaft, also der zeitlichen Distanz zum Ereignis, darstellt. Sie berechnen für Zeiträume von zwei und sechs Jahren eine Zunahme des Datierungsfehlers um 0,15 Tage pro verstrichenem Tag, also um einen Tag pro Woche ( $0,15 \times 7 = 1,05$ , siehe Abb. 2.6). Linton (1975) kommt bei einem retrospektiven Intervall von vier Wochen auf 0,22 Tage pro verstrichenem Tag; und Thompson (1982) auf 0,16 Tage. Ein vier Monate zurückliegendes Ereignis würde demnach mit einem Fehler von etwa 16 bis 18 Tagen datiert werden. Baddeley, Lewis und Nemo-Smith (1978) und Thompson (1982) weisen einen schwach kurvilinearen Abfall des Anteils korrekter Daten über 10 bis 15 Wochen nach: In den ersten 40 Tagen ist der Abfall sehr steil und wird dann zunehmend flacher. Auch die Größe der abweichenden Datierungen nimmt über 10 und 15 Wochen hin zuerst steiler, später schwächer zu; über 30 Monate hingegen steigt sie linear an.

Mehrere Studien belegen auch, dass die Richtung des Fehldatierens sich mit der Länge der Rückschau ändert und die Neigung zum Vordatieren mit der Länge steigt: „Ältere“ Ereignisse werden eher nach vorn verlegt und „jüngere“ eher nach hinten (Baddeley u.a., 1978; Brown

u.a., 1985; Burt u.a., 2001). Auch Thompson (1996) findet für retrospektive Distanzen von bis zu 10 Wochen, bis zu 15 Wochen und bis zu 30 Monaten kein einseitiges Teleskopieren, sondern dass weiter zurückliegende Ereignisse nach später und kürzer zurückliegende Ereignisse nach früher fehldatiert wurden. Dies kann auch als Neigung beschrieben werden, Ereignisse weg vom Rand und hin zur Mitte des Zeitraums zu datieren; die Gründe hierfür wurden in Abschnitt 2.5.3 besprochen.

Generell läßt also sowohl die Erinnerungsgüte als auch die Güte des Datierens mit zunehmender Distanz zwischen Erleben und Rekonstruieren nach, wobei die Geschwindigkeit des Nachlassens von der Art der zu erinnernden Inhalte abhängt. Die genaue Form der Kurve fällt ebenfalls unterschiedlich aus: Für die Größe der Datierungsfehler wurde eine lineare Verschlechterung belegt, ansonsten scheint eine Gütefunktion den Zusammenhang am besten zu beschreiben.

Bei der Betrachtung von Vergessenskurven, wie den berichteten, wird oft davon ausgegangen, dass *ceteris paribus* außer der retrospektiven Distanz keine Veränderungen an den zu erinnernden Ereignissen bzw. den Umständen der Rekonstruktion auftreten. So behandelte die Studie von Rubin und Baddeley (1989) jedes Kolloquium, das die Teilnehmer zu datieren hatten, gleich; sie unterschieden nicht nach (beispielsweise) Kolloquien namhafterer Vortragender oder Kolloquien, die direkt vor oder nach Weihnachten stattfanden und sich durch diese Nähe zu einem Grenzstein leichter erinnern oder datieren lassen sollten. Beim Rekonstruieren autobiografischer Information unter Alltagsbedingungen oder in Befragungen hingegen ist die Vergleichbarkeit aller anderen Umstände für jedes zu rekonstruierende Ereignis und jeden Punkt der retrospektiven Periode selten gegeben. Im Folgenden möchte ich einige Faktoren besprechen, die mit der retrospektiven Distanz einhergehen und die Vergessens- bzw. Fehldatierungskurve in ihrer Neigung oder ihrer Form ebenfalls beeinflussen. Dies sind vor allem *Alter bei Erinnern* (siehe Abschnitt 2.5.2), *Alter bei Erleben* und *historische Veränderungen*.

### 2.6.3 Die retrospektive Distanz und mit ihr einhergehende Faktoren

Unterschiedlich retrospektive Distanzen können zum einen damit einhergehen, dass die Erinnernden unterschiedlich alt sind: Eine 50 Jahre zurückliegende Lehre kann fast nur von einer Person erinnert werden, die mindestens Mitte 60 ist. Die Rolle des Alters beim Erinnern spielt eine Rolle für die Erinnerungsgüte, wie ich in Abschnitt 2.5.2 ausführlich dargelegt habe, und kann so den Zusammenhang zwischen retrospektiver Distanz und Erinnerungsgüte mit beeinflussen.

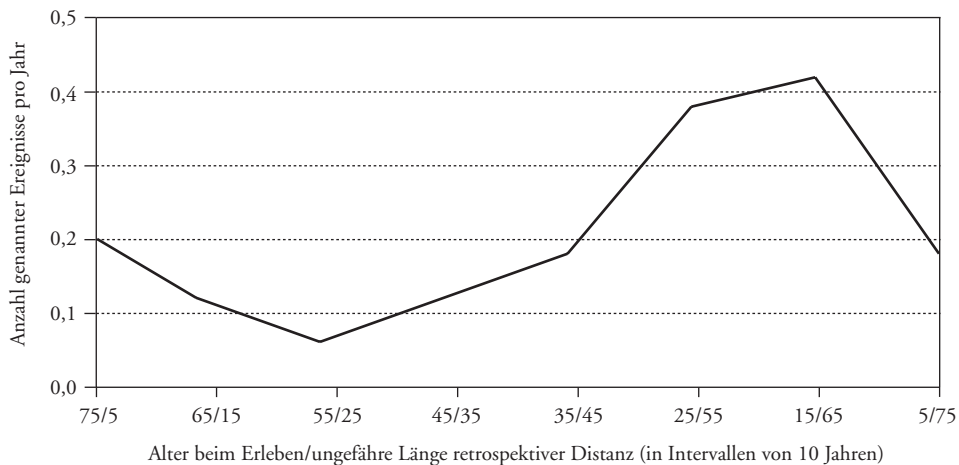
Weiterhin können unterschiedliche retrospektive Distanzen mit Zeiträumen zusammenfallen, in denen die Erinnernden ein bestimmtes Alter hatten. Dass dies für die Güte der Erinnerung relevant ist, wurde in Studien belegt, in denen die Studienteilnehmer entweder nach den wichtigsten Ereignissen ihres Lebens gefragt werden oder nach den Ereignissen, die ihnen spontan auf ein Schlüsselwort hin einfallen (Crovitz & Schiffman, 1974). Man geht davon aus, dass die genannten Ereignisse die höchste Verfügbarkeit bzw. Zugänglichkeit besitzen und deshalb mit größerer Wahrscheinlichkeit behalten und auch korrekt erinnert werden können. In solchen Studien findet sich dabei konsistent, dass die deutliche Mehrheit aller genannten Erinnerungen aus der Zeit unmittelbar vor der Untersuchung stammt. Die Verteilung der übrigen Erinnerungen ist dann aber nicht monoton abfallend, vielmehr wurden zwei

Effekte des Alters beim Erleben einer Erinnerung belegt: die „frühkindliche Amnesie“ und das „Reminiszenzphänomen“. Frühkindliche Amnesie bedeutet, dass aus den ersten etwa fünf Lebensjahren fast keine Erinnerungen verfügbar und diese in der Regel vage und arm an Details und Kontexten sind (Fitzgerald, 1996; Rubin, 1982; Rubin u.a., 1986; Wetzler & Sweeney, 1986). Dies lässt sich darauf zurückführen, dass elaborierte Strukturen im autobiografischen Gedächtnisbestand erst ab einem gewissen Alter zu erwarten sind.

Reminiszenzphänomene traten bei Personen über 40 Jahre auf: Die Teilnehmer nennen unerwartet häufig Erinnerungen aus der Zeit, als sie 15 bis 30 Jahre alt waren. Die Zeit der Adoleszenz und des frühen Erwachsenenalters ist die ausschlaggebende Phase für die Entwicklung des Selbst und der stabilen Identität einer Person (Habermas & Bluck, 2000; Rubin u.a., 1986). Wenn die Studienteilnehmer nach persönlich bedeutsamen oder besonders lebhaften Erinnerungen gefragt wurden, war das Reminiszenzphänomen noch ausgeprägter als bei der Verwendung der Schlüsselwortmethode, und die Häufung kurz zurückliegender Ereignisse blieb aus oder war schwächer. Abbildung 2.7 zeigt eine Verfügbarkeitskurve aus einer Studie, in der die über 70 Jahre alten Teilnehmer zu ihren wichtigsten Erinnerungen befragt wurden. Ihr Verlauf lässt sich als Ergebnis dreier Faktoren erklären: Zunächst werden vergleichsweise viele Erinnerungen aus den Jahren direkt vor der Befragung angegeben (Effekt der retrospektiven Distanz). Eine ausgeprägte Häufung findet sich bei Alter 15 bis 25 (Reminiszenzphänomen), und aus der Zeit, als die Befragten unter fünf Jahre alt waren, werden überhaupt keine Ereignisse berichtet (frühkindliche Amnesie). Die Validität dieser Erinnerungen wurde hier allerdings nicht untersucht.

Andererseits halten es etwa Fromholt und Larsen (1994) für möglich, dass die seltenere Nennung von Erinnerungen aus dem mittleren Alter auch die historisch bedingte Kontinuität und Stabilität der mittleren Lebensjahre dieser speziellen Kohorten widerspiegelt, die nach der

Abbildung 2.7: Zugänglichkeit von Erinnerungen über die Lebensspanne



Die Verteilung der wichtigsten Ereignisse eines Lebens über sieben Dekaden der Lebensspanne bzw. sieben Dekaden retrospektiver Distanz. Die Kurve ergibt sich aus Vergessen, Reminiszenzphänomen und frühkindlicher Amnesie.

Quelle: Fromholt & Larsen, 1994.

frühen Ausbildung, Eheschließung, Familiengründung und beruflichen Etablierung erst wieder im höheren Alter durch Verrentung, Verwitwung oder nachlassende Gesundheit unterbrochen wird. Wenn historische Einflüsse wie Kriege und Wirtschaftskrisen alle Altersgruppen ohne Ausnahme betreffen, können auch die mittleren Jahre durch zahlreiche Brüche und Übergänge charakterisiert sein (Brückner, 1993). Auch durch den Umstand, dass sich jüngere Kohorten familiär und beruflich tendenziell immer später etablieren, sowie durch die erhöhte berufliche Mobilität und steigende Scheidungsraten mögen auch die mittleren Jahre jüngerer Kohorten wieder verstärkt von Veränderungen und Übergängen in den sozialen Rollen gekennzeichnet sein. Insofern können, zumindest bei Querschnittstudien, die Vergessens-, Fehldatierungs- und Verteilungsfunktionen auch für historische Veränderungen stehen und Effekte als Perioden- oder Kohorteneffekte interpretiert werden.

Eine längere retrospektive Distanz steht also mit schlechterem Erinnern im Zusammenhang und macht damit das Auftreten von Abweichungen wahrscheinlicher. Das geschieht auf vielen Wegen, unter anderem durch die Zunahme an inferenziell rekonstruktivem Erinnern, durch das Verblässen von Erinnerungsdetails, durch das Alter bei Erinnern, durch das Alter beim Erleben, durch historische Veränderungen, durch Veränderungen der aktuellen den Abruf anleitenden Schemata und durch Veränderungen der Repräsentationsstrukturen. Diese Überlegungen verdeutlichen nachdrücklich, dass eine Vergessenskurve autobiografischer Inhalte in der realen Welt stets das kumulative oder interaktive Ergebnis mehrerer Prozesse und Faktoren ist, deren Einfluss sich durch und mit der verstrichenen Zeit ändert.

In den vorigen Abschnitten habe ich die Mechanismen des autobiografischen Erinnerns, die durch sie verursachten Unterschiede zwischen Rekonstruktionen und den tatsächlichen Sachverhalten beschrieben. Vor diesem Hintergrund habe ich betrachtet, welche Information besonders anfällig ist und den Zusammenhang mit der Länge der retrospektiven Rückschau beleuchtet. Im letzten Abschnitt werde ich die zentralen Befunde zusammenfassen und in Bezug zu retrospektiven Befragungen setzen.

## **2.7 Rekonstruktionsfehler beim autobiografischen Erinnern – Ein Rahmen für Fehler in retrospektiven Befragungen?**

Die kognitionspsychologischen Theorien zum autobiografischen Erinnern bieten einen ausgezeichneten Hintergrund für die Generierung von spezifischen Hypothesen über Erinnerungs- und Datierungsfehler in retrospektiven Befragungen, da sie die Mechanismen des Erinnerns herausarbeiten und die Bedingungen untersuchen, unter denen diese Mechanismen verschiedene Arten von abweichenden und zeitlich instabilen Rekonstruktionen verursachen. Sie beschreiben das Erinnern autobiografischer Information als einen selektiven, interpretativen und dynamischen Prozess, durch den Rekonstruktionen nicht immer zeitstabil ausfallen und die erinnerte von der tatsächlich erlebten Autobiografie sowie spätere von früheren Rekonstruktionen auf charakteristische Weise abweichen. Durch selektives Vergessen und Umdeuten wird in der Rückschau Folgerichtigkeit im Sinne der Ziele und Entwicklungslogiken des aktuellen Selbst hergestellt. Außerdem wird eine narrative Dramaturgie geschaffen, innerhalb derer Ereignisse – oder deren Ausbleiben – als Folge der Vergangenheit und Vorgeschichte der Gegenwart begründet werden. Dadurch sind Erinnerungen vereinfachte, kohärentere Versionen der tatsächlichen Sachverhalte. Diese Versionen stehen auch stärker im Einklang mit nor-



mativen Konventionen und Erwartungen. Im Allgemeinen führt das auch dazu, dass negativ empfundene Ereignisse bevorzugt vergessen oder positiver umgedeutet werden. Fehler beim Datieren autobiografischer Ereignisse sind sehr verbreitet, in der Regel sind sie aber unsystematisch und betragen nur wenige Zeiteinheiten.

In retrospektiven Befragungen, wenn der Befragte die von ihm verlangte Information aus den Repräsentationen konstruiert, werden sich derlei Tendenzen ebenfalls bemerkbar machen und sich auf Datenreliabilität und -validität auswirken. Nicht alle Informationen werden davon gleichermaßen betroffen sein. Zusammenfassend kann man sagen, dass Information auch über längere retrospektive Distanzen zeitstabiler und korrekter erinnert und datiert wird, wenn sie:

- den Repräsentationen im autobiografischen Gedächtnis entspricht, wenn also Lebensabschnitte, generische Ereignisse, Minigeschichten usw. erinnert werden sollen,
- sich in die schematischen Selbststrukturen der aktuellen Erinnerungssituation einfügt,
- direkt replikativ erinnert werden kann,
- reichhaltig verknüpft ist,
- häufig geübt und gebahnt wurde,
- eine zentrale Position innerhalb der schematischen Strukturen und Lebensgeschichten einnimmt,
- viele Details als inhaltliche und temporale Abruf-Hinweisreize aufweist<sup>9</sup>.

Dies trifft vor allem auf emotional aufgeladene und einzigartige Ereignisse zu, die Folgen für das Leben des Befragten hatten. Für solche Information kann man also auch in retrospektiven Befragungen gute Qualität erwarten. Am wenigsten ist es der Fall für Kalenderdaten und andere Fakten ohne Erlebnisgrundlage (z.B. Einkommensangaben oder Schulnoten).

Abweichungen treten generell durch die retrospektive Rekonstruktion auf. Ob und wie sie darüber hinaus mit der Länge der retrospektiven Distanz zunehmen oder sich in ihrer Art und Richtung ändern, hängt von den zu Erinnernden Inhalten ab. Besonders für die zentralen Bestandteile der Lebensgeschichte ist aufgrund der hohen Organisiertheit und Verknüpftheit sowie durch die hohe Übungsfrequenz eine hohe zeitliche Stabilität und Korrektheit der Erinnerung anzunehmen. Unterschiedliche retrospektive Distanzen können außerdem mit Unterschieden in mehreren erinnerungsrelevanten Bedingungen einhergehen: Je nachdem, wie eine Studie aufgebaut ist, müssen das Alter beim Erinnern, das Alter beim Erleben sowie die Kohortenzugehörigkeit und historische Veränderungen berücksichtigt werden. Für retrospektive Befragungsdaten, die lange Zeiträume abdecken, muss also differenziert ermittelt werden, inwieweit die Angaben zu den sehr weit zurückliegenden Ereignissen wirklich deutlich schlechter sind als die zur jüngeren Vergangenheit und welcher der relevanten Faktoren an der Form der Vergessenskurve beteiligt ist.

Ausmaß und Art der Abweichungen, die beim Rekonstruieren auftreten, hängen davon ab, welche konkreten kognitiven Schritte beteiligt sind und welches die Rahmenbedingungen der Rekonstruktion sind. Besonders in Abschnitt 2.5.3 ist im Rahmen der Telescoping-Kontroverse deutlich geworden, dass die Erklärung einer bestimmten Art von Abweichung

<sup>9</sup> Diese Faktoren sind nicht unabhängig voneinander: Detaillierte und vielfach verknüpfte Erinnerungen, die als Bestandteil der Strukturen repräsentiert sind, werden häufiger abgerufen. Dadurch verblässen die Details und Verknüpfungen langsamer. Solche Informationen werden auch am ehesten zeitstabil sein, denn als zentrale Bestandteile der Struktur sind sie änderungsresistent.

davon abhängt, welche Annahmen über das Zustandekommen der Antwort bestehen. Deshalb ist es bei der Erklärung von Abweichungen wichtig, sehr genau darauf zu achten, welche Fragen gestellt und wie Antworten festgehalten werden.

Im nächsten Kapitel möchte ich darlegen, wie Fehler in standardisierten retrospektiven Längsschnittbefragungen als Ergebnis der Erinnerungsprozesse erklärt werden können, die unter den speziellen Rekonstruktionsbedingungen der Befragungssituation ablaufen. Ich werde hierzu relevante Konzepte aus den Theorien des Messens und Datengenerierens sowie der befragungsmethodologischen Literatur aufgreifen, die Fehler beschreiben, Datengüte quantifizieren und als Folge von Datengenerierungsprozessen erklären. In diesen Kontext werde ich die Rolle von Gedächtnisprozessen einordnen.



## Kapitel 3

### Datengenerierung und Datenqualität – Konzepte, Begriffe und Befunde

In diesem Kapitel möchte ich darstellen, wie Fehler in standardisierten retrospektiven Befragungen als Ergebnis von Erinnerungsprozessen in der Befragungssituation erklärt werden können. Zuerst werde ich eine Klärung und Elaborierung des Fehlerbegriffs für Ereignisdaten vornehmen. Danach werde ich zum einen herausarbeiten, wie man Fehler aussagekräftig, umfassend und differenziert beschreiben kann, indem man ihre Verbreitung, ihr Ausmaß, ihre Richtung und ihre Abhängigkeit vom tatsächlichen Sachverhalt betrachtet. Des Weiteren werde ich Konzepte aus der Befragungsmethodologie vorstellen, die einen konzeptuellen Rahmen für die Prozesse der Daten- und Fehlerproduktion in standardisierten Längsschnittbefragungen bieten, und die Rolle von Erinnerungsprozessen in denselben einordnen. Danach werde ich einen Überblick über die empirischen Befunde zur Güte retrospektiv erhobener Ereignisdaten in der Lebensverlaufsorschung geben. Zuletzt werde ich auf diesem Hintergrund die Studie vorstellen, auf der die empirischen Analysen im folgenden Kapitel beruhen und die Bedeutung ihrer Konzepte und Erhebungstechniken für die Erinnerungsprozesse der Befragten (und damit für die Datenfehler) erörtern.

#### 3.1 Daten, Realität und Fehler

##### 3.1.1 Fehler in Ereignisdaten

Daten sind Informationen über Eigenschaften von Objekten oder Merkmalsträgern, die in standardisierter Form – nämlich als numerische Werte einer Variablen – Aspekte der Realität abbilden bzw. repräsentieren. Variablen können mit ihren numerischen Werten sowohl qualitativ unterschiedliche Zustände (erwerbstätig und nichterwerbstätig oder männlich und weiblich) oder quantitative Unterscheidungen (Alter, Intelligenz oder Dauer der Betriebszugehörigkeit) anzeigen. Im ersten Fall spricht man auch von kategorialen Daten.

Die Annahmen über das Verhältnis der in den Daten repräsentierten Information zur Realität bestimmen den Rahmen, in dem wir von Datenfehlern sprechen. In den Sozialwissenschaften verbreitet ist eine (implizite oder explizite) Position des *kritischen Realismus*. Dieser nimmt an, dass sozialwissenschaftliche Konzepte Ausschnitte oder Aspekte einer unabhängigen Realität begrifflich fassen. Das bedeutet, dass es gemäß der Definition des Sozialwissenschaftlers einen wahren Zivilstatus einer Person zu einem bestimmten Zeitpunkt, die wahre Dauer einer Erwerbstätigkeit und einen wahren Zeitpunkt einer Eheschließung tatsächlich gibt. Abweichungen zwischen tatsächlichen Sachverhalten und ihren Repräsentatio-

nen in den Daten werden als Fehler bezeichnet und auf die Prozesse zurückgeführt, die zwischen Realität und Repräsentation vermitteln<sup>1</sup>.

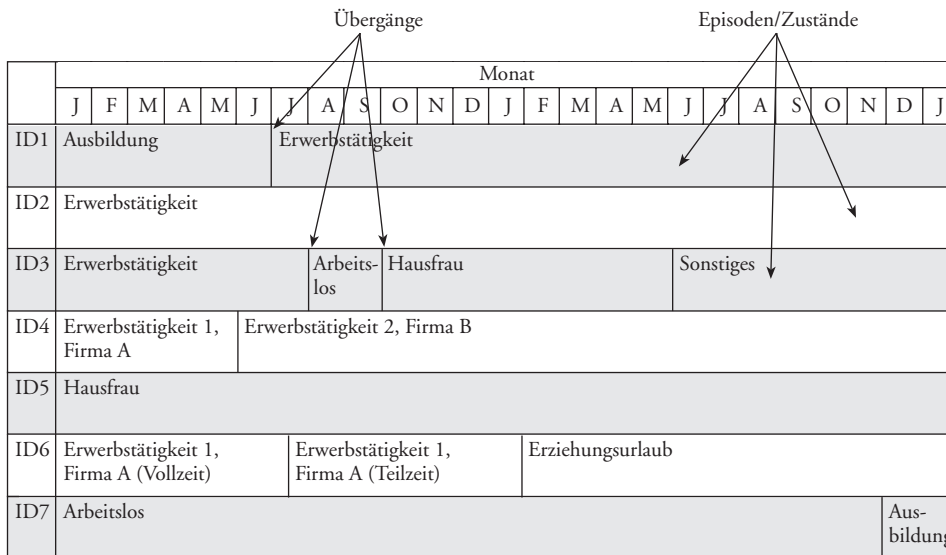
Fehler sind unvermeidlich. Nur in einer perfekten Welt mit unfehlbaren Datengenerierungsprozessen würden sich Repräsentation und tatsächlicher Sachverhalt immer exakt entsprechen. Folglich sind Prozesse der Datengenerierung immer auch Prozesse der Datenfehlergenerierung. Da Sozialwissenschaftler also in jedem Fall mit fehlerbehafteten Daten zu tun haben werden, müssen sie versuchen, den Fehlereinfluss durch eine geeignete Untersuchungsplanung im Vorfeld zu minimieren und ihn nach der Datenerhebung bei der Auswertung und Interpretation zu berücksichtigen. Vorbeugende Maßnahmen gegen das Auftreten von Fehlern lassen sich nur konzipieren und in ihrer Kosteneffektivität einschätzen, wenn man den Mechanismus der Fehlerentstehung kennt. Und derjenige, der Fehler bei der Befundinterpretation berücksichtigen will, benötigt dieses Verständnis, um zu beurteilen, mit welchen Fehlern er in einem konkreten Datensatz aufgrund dessen Entstehungsbedingungen rechnen muss oder auf welche Personen, Daten und Umstände er die Fehlerbefunde an einer konkreten Datenmenge verallgemeinern kann. Hierzu müssen Fehler aussagekräftig beschrieben und als Folge der Datengenerierungsprozesse erklärt werden, um das Verhältnis von Realität zu Datumswert erkennbar zu machen.

Ereignisdaten verfolgen ihre Merkmalsträger – meist Individuen – über die Zeit hinweg und halten sowohl inhaltliche als auch zeitliche Informationen über die interessierenden Zeiträume fest. Hierzu werden alle Ereignisse eines bestimmten Typs erfasst, etwa jede Erwerbstätigkeit oder jede Erkrankung. Danach werden für jedes einzelne Ereignis inhaltliche und zeitliche Detailinformationen, etwa Dauer, Anfang- und Enddaten, erhoben. Zeitlich ausgedehnte Ereignisse bezeichnet man als Episoden. Sie werden von zeitlich nicht ausgedehnten Übergängen begrenzt, an denen das Individuum eine Episode verlässt und in einen anderen Zustand überwechselt. Beispielsweise werden Wohnepisoden durch Umzüge begrenzt, Krankheitsepisoden durch Genesung und Erwerbsepisoden durch die Beendigung des Erwerbsverhältnisses. Entweder werden nur Ereignisse eines Typs – etwa alle Krankenhausaufenthalte oder Erwerbstätigkeiten – erfasst, beschrieben und datiert; wenn es sich um zeitlich ausgedehnte Episoden handelt, erfolgt der Wechsel immer von einer Episode in eine „Nicht-episode“, und die Variable hat nur zwei Ausprägungen, etwa „erwerbstätig“ und „nicht erwerbstätig“. Oder es wird eine endliche Menge an sich gegenseitig ausschließenden Zuständen definiert – zum Beispiel „ledig, verheiratet, geschieden, verwitwet“ oder „erwerbstätig, arbeitslos, in Ausbildung, Hausfrau und Sonstiges“. Die interessierenden Zeiträume können dann als lückenlose „Wanderung“ durch diesen Zustandsraum beschrieben werden. So entstehen Ereignissequenzen wie die in Abbildung 3.1 dargestellten (fiktiven) Erwerbsverläufe, wie sie auch in der „deutschen Lebensverlaufsstudie“ (LV) erhoben werden.

In solchen Daten ist eine enorme Fülle inhaltlicher und zeitlicher Information repräsentiert, anhand derer die abgedeckten Zeiträume durch eine Reihe von Verlaufseigenschaften

<sup>1</sup> Die Debatte über Realität und ihre Repräsentation ist keineswegs abschließend geklärt. Fehlerforscher und Befragungsmethodiker legen aber oft eine pragmatische Haltung an den Tag. So meint Kruskal (1991): „There is a widespread desire to cling to faith in an objective reality out there (...) If there is no true value lurking in the platonic underbrush, how can we talk sensibly about errors?“ Und Trochim (2000): „In the end, I tend to turn pragmatist on these matters. Philosophers have been debating these issues for thousands of years, and there is every reason to believe that they will continue to debate them for thousands of years more. Those of us who are practising scientists should check in on this debate from time to time (perhaps every hundred years or so would be right).“

Abbildung 3.1: Lebensgeschichtliche Ereignisdaten (schematische Darstellung)



ten charakterisiert werden können (Caspi u.a., 1996; Mayer & Huinink, 1990; Rohwer, 1994):

- *Inzidenz und Frequenz* geben an, ob und wenn ja wie viele Ereignisse eines Typs im abgedeckten Zeitraum auftreten;
- das *Timing* gibt an, zu welchem Zeitpunkt in der historischen oder Lebenszeit Episoden oder Ereignisse auftreten, bzw. welche Dauer bis zum Eintreten eines bestimmten Ereignisses verstreicht;
- die *Dauer* oder Dauerverteilung gibt an, wie lange Episoden andauern bzw. wie viel Zeit jemand in einem oder allen zur Verfügung stehenden Zuständen verbringt;
- die *Reihenfolge* beschreibt, wie Ereignisse oder Zustände aufeinander folgen.

Fehler allgemein und Gedächtnisfehler im Besonderen können also in ihrem Einfluss auf inhaltliche und zeitliche Informationsaspekte der Episoden, der Übergänge zwischen ihnen und der gesamten Sequenzen beschrieben und erklärt werden.

### 3.1.2 Fehler beschreiben

Der erste Schritt bei der Erklärung von Datenfehlern als Ergebnis von Erinnerungsprozessen besteht darin, den Fehler aussagekräftig zu beschreiben. Dabei interessieren zum einen die individuelle Abweichung *einer* Angabe *einer* Person vom tatsächlichen Sachverhalt und zum anderen, wie häufig individuelle Abweichungen in einer Datenmenge auftreten oder welche Art von Abweichungen häufiger als andere auftritt.

### *Irrtum und Fehldiagnose*

Woolgar (1983) bezeichnet Berichte oder Daten als von der Realität unterdeterminiert: Durch den Fehler kann ein und derselbe Datumswert für verschiedene reale Sachverhalte stehen und umgekehrt kann derselbe reale Sachverhalt durch unterschiedliche Datumswerte repräsentiert sein. Fehler, also Abweichungen zwischen Realität und Datenabbild, können in beiderlei Hinsicht beschrieben werden: Ausgehend von der Realität möchte man wissen, welcher tatsächliche Sachverhalt welchen Datumswert zur Folge hat. Umgekehrt, ausgehend von den Daten, möchte man wissen, beim Vorliegen welchen Datumswerts mit dem Vorliegen welcher Realität gerechnet werden muss. Diese beiden „Richtungen“ des Fehlers können als Irrtum und Fehldiagnosen charakterisiert werden. Ein Beispiel: Beim Wiegen einer 75 kg schweren Person interessiert mich, mit welchen Gewichtsanzeigen ich rechnen muss – wie oft wird die Anzeige fehlerhaft sein und welche Werte werden stattdessen angezeigt? Andererseits interessiert mich, wenn eine Waage 75 kg anzeigt, ob und wie oft diese Anzeige tatsächlich für 75 kg schwere Personen steht und wie schwer diese Personen tatsächlich waren. Um zu Erkenntnissen über den Datengenerierungsprozess zu gelangen, der ausgehend von bestimmten Sachverhalten bestimmte Datenwerte ergibt, ist vor allem die Information über den Irrtum interessant.

### *Vier Aspekte des Fehlers*

Irrtum und Fehldiagnosen können hinsichtlich vier Eigenschaften beschrieben werden: die Verbreitung des Fehlers, das Ausmaß des Fehlers, die Art oder Richtung des Fehlers und die Abhängigkeit bzw. Unabhängigkeit dieser Eigenschaften vom realen Sachverhalt.

Die *Verbreitung* des Fehlers beschreibt, ob in einem individuellen Fall überhaupt Abweichungen vom realen Sachverhalt auftreten und bei wie vielen Fällen innerhalb einer Menge von Daten dies der Fall ist. Die Informationen einer Waage, die von 100 Personen für 90 das korrekte Gewicht anzeigt, ist von der Verbreitung her weniger fehlerhaft, als wenn dies nur für 75 Personen zutrifft; das Irrtumsrisiko ist hier höher. Umgekehrt ist die Verbreitung von Fehldiagnosen höher, wenn 100 identische Gewichtsanzeigen nur für 75 Personen zutreffen, als wenn 90 Personen zu Recht in der jeweiligen Gewichtskategorie eingeordnet werden. Auch für kategoriale Daten kann man die Fehlerverbreitung beschreiben: Wenn 20 von 100 Personen ihren Zivilstand falsch angeben, sind die Daten von der Verbreitung her fehlerhafter und das Irrtumsrisiko höher, als wenn nur 5 Personen falsch antworten.

Das *Ausmaß bzw. die Größe* des Fehlers beschreibt, wie weit der Datumswert bei einem Individuum vom tatsächlichen Sachverhalt abweicht bzw. wie weit die abweichenden Angaben um die tatsächlichen Sachverhalte streuen. Je besser Realität und Datum übereinstimmen, desto geringer ist die Bandbreite von Datumswerten, die für denselben tatsächlichen Sachverhalt stehen kann, und desto geringer ist die Bandbreite der tatsächlichen Werte, die sich hinter einem bestimmten Datum verbergen können. Eine Waage, die für 60 kg schwere Menschen Gewichtsanzeigen zwischen 59,5 und 60,5 kg ausgibt, ist von der Irrtumsgröße her weniger fehlerhaft als eine, die Gewichte zwischen 50 und 70 kg ausgibt. Umgekehrt ist eine Waage, die 60 kg für Menschen von 59,5 bis 60,5 kg anzeigt, weniger fehleranfällig als eine, die auch für Personen mit 50 oder 70 kg 60 kg anzeigen würde.

Für kategoriale Daten ist das Ausmaß der Abweichungen nicht direkt erkennbar. Handelt es sich um eine „größere“ Abweichung, wenn eine verheiratete Person angibt, ledig zu sein, als

wenn sie angibt, geschieden zu sein? Dies kann allenfalls aufgrund der Absichten des erhebenden oder auswertenden Sozialwissenschaftlers entschieden werden, wenn man die Fehlergröße in dem Sinne versteht, welche Arten von Abweichungen gravierendere Folgen für die Schlussfolgerungen haben<sup>2</sup>.

Die *Richtung des Fehlers* beschreibt, ob eine abweichende Angabe größer oder kleiner, höher oder niedriger als der tatsächliche Sachverhalt ausfällt, bzw. ob in einer Datenmenge die abweichenden Angaben zu gleichen Teilen höher und niedriger ausfallen oder Fehler in eine Richtung häufiger auftreten. Die Information einer Waage, die das Gewicht der sich wiegenden Person gleich häufig und in gleichem Ausmaß über- und unterschätzt, ist auf der Irrtumsseite mit einem symmetrischen oder ungerichteten Fehler behaftet; eine Waage, die in den meisten Fällen ein zu hohes Gewicht anzeigt hingegen mit einem gerichteten Fehler. Auf der Seite der Fehldiagnosen ist die Information einer Waage mit einem gerichteten Fehler bei den Fehldiagnosen behaftet, wenn die Gewichtskategorien vor allem Personen beinhalten, die eigentlich weniger oder mehr wiegen. Für kategoriale Daten ist die Fehlerrichtung weniger direkt aussagekräftig. Es handelt sich allerdings um einen anders gelagerten Fall, wenn auf alle, die eine falsche Angabe über ihren Zivilstand machen, stattdessen „ledig“ angeben, als wenn sie alle „verheiratet“ angeben oder sich die Antworten zu gleichen Teilen auf die Kategorien „ledig“, „verheiratet“ und „geschieden“ verteilen.

Wenn der Fehler ungerichtet ist und die abweichenden Datumswerte symmetrisch um den tatsächlichen Sachverhalt streuen, können sich die Abweichungen über alle Beobachtungen hinweg gegeneinander aufheben, sodass die Daten zumindest für aggregierte Stichprobenkennwerte wieder eine korrekte Schätzung ergeben. Wenn hingegen die Abweichungen in eine bestimmte Richtung häufiger vorkommen oder größer ausfallen, ist dies nicht mehr möglich und aggregierte Information wird verzerrt abgebildet. Insofern stellt ein gerichteter Fehler das größere Problem dar.

Zuletzt muss geprüft werden, ob der Fehler *in Abhängigkeit vom tatsächlichen Sachverhalt* auftritt. Dies ist dann der Fall, wenn unterschiedliche tatsächliche Sachverhalte mit Fehlern von unterschiedlicher Verbreitung, unterschiedlicher Größe oder unterschiedlicher Richtung einhergehen. Die Fehler einer Waage, die für 60 kg schwere Menschen sehr häufig abweichende und immer zu hohe Werte anzeigt, während sie für Menschen mit 70 kg deutlich seltener abweichende und eher zu niedrige Gewichtsdaten liefert, wären also vom tatsächlichen Sachverhalt abhängig. Umgekehrt wären die Fehldiagnosen abhängig vom tatsächlichen Sachverhalt, wenn 60 kg ausschließlich für Personen angezeigt wird, die auch wirklich 60 kg schwer sind, während 70 kg fast nur für Personen angezeigt wird, die eigentlich weniger wiegen. Für kategoriale Daten liegen abhängige Fehler vor, wenn alle Verheira-

---

<sup>2</sup> Wenn beispielsweise die aktuell Verheirateten den aktuell Nichtverheirateten gegenübergestellt werden sollen, wäre es nicht so gravierend, wenn eine ledige Person ihren Zivilstand mit „geschieden“ angibt, da sie dennoch unter den Nichtverheirateten geführt werden würde. Wenn sie hingegen angibt, „verheiratet“ zu sein, würde sie fälschlicherweise der Gruppe der Verheirateten zugeschlagen werden. Hier könnte man also sagen, dass es einen größeren, das heißt gravierenderen Fehler darstellt, wenn eine ledige Person sich als verheiratet einordnet als wenn sie sich als geschieden einordnet. Sollen hingegen alle jemals Verheirateten (egal ob sie mittlerweile geschieden sind) den Nie-Verheirateten gegenüber gestellt werden, würde die ledige Person die Gruppe in beiden Fällen wechseln, was beide Fehler gleich „groß“ erscheinen lässt. Die Abgrenzung zum Aspekt der Fehlerrichtung (siehe nächster Punkt) wird dann allerdings schwierig.



reten ihren Zivilstand auf eine Frage hin korrekt angeben, während Ledige sich sehr häufig falsch einordnen, oder sich in die Gruppe der getrennt Lebenden zwar alle getrennt Lebende korrekt einordnen, dazu aber noch einmal genau so viele eigentlich Geschiedene kommen.

Praktisch kann Abhängigkeit der Fehler ein Hinweis darauf sein, dass die Skala in bestimmten Bereichen nicht zur Merkmalerfassung taugt bzw. dass eine der Kategorienausprägungen nicht geeignet ist, die Objekte zu charakterisieren. Eine Waage, deren Mechanik das korrekte Klassifizieren von Gewicht nur innerhalb einer Bandbreite von 15 kg bis 140 kg erlaubt, stellt für Kleinkinder oder stark Übergewichtige ein ungeeignetes Instrument zur Gewichtserfassung dar. Ebenso könnte es sein, dass für bestimmte Populationen das Konzept der „Ledigen“ keine tatsächlich existierende, abgrenzbare Gruppe beschreibt, das Konzept der „Verheirateten“ hingegen schon.

Die beschriebenen Fehleraspekte sind weitgehend unabhängig voneinander<sup>3</sup>. Wenn Fehler untersucht und beschrieben werden sollen, muss also angegeben werden, was genau betrachtet wird. Welcher Informationsaspekt der Realität soll in den Daten abgebildet werden? Für welchen in den Daten abgebildeten Informationsaspekt soll der Prozess seiner Abbildung und seine Fehlerhaftigkeit betrachtet werden? Reicht eine Übereinstimmung aggregierter Stichprobenkennwerte oder sollte jede Angabe auf Individualebene korrekt sein? Stehen Irrtümer oder Fehldiagnosen im Vordergrund? Welche der vier Aspekte des Fehlers werden untersucht?

Um darüber hinaus zu handlungsleitenden Erkenntnissen zur Vorbeugung und Abschätzung von Datenfehlern zu gelangen, muss das Zustandekommen der umfassend und differenziert beschriebenen Datenfehler aus den Eigentümlichkeiten des Datengenerierungsprozesses unter Rückgriff auf relevantes Wissen über dessen Verlauf und Umstände erklärt werden. Dazu müssen wir jeweils für die konkrete Datengenerierungsprozedur bzw. -situation modellhaft spezifizieren, welche Prozesse zwischen der Realität und Daten vermitteln und welche Faktoren auf der Seite der zu erhebenden Information, der befragten Personen und der Erinnerungssituation die Übereinstimmungen von Realität und Datum beeinflussen. Im folgenden Abschnitt werde ich Konzepte und Modelle aus der Fehlerliteratur vorstellen, die sich mit diesen Fragen befassen, und erörtern, inwieweit sie zur Beschreibung der Datengenerierung in retrospektiven Befragungen herangezogen werden können.

### 3.2 Daten- und Datenfehlergenerierung in retrospektiven Befragungen

#### 3.2.1 Die Konstruktion sozialwissenschaftlicher Daten als Variablenwerte

Rohwer und Pötter (2002) beschreiben die sozialwissenschaftliche Datenproduktion als einen mehrstufigen Konstruktionsprozess: „Daten sind Werte statistischer Variablen. Man beginnt mit der Definition einer statistischen Variablen (...) und versucht dann, für die Mitglieder

---

<sup>3</sup> Dies gilt mit der Einschränkung, dass Fehler natürlich nur dort, wo sie überhaupt und in einer gewissen Anzahl auftreten, als Irrtum und Fehldiagnosen betrachtet bzw. hinsichtlich ihrer Größe, Richtung und Abhängigkeit eingeschätzt werden können.

Abbildung 3.2: Der Prozess sozialwissenschaftlicher Datenkonstruktion

---

Schritt 1	Definition der Objekte
Schritt 2	Definition des Merkmals- oder Zustandsraums
Schritt 3	(a) Ermittlung der Objekte (Stichprobenziehung) (b) Ermittlung der Merkmals- oder Zustandswerte für die Objekte (Messung)

---

Quelle: Rohwer & Pötter, 2002.

(...) Werte der Variablen zu ermitteln.“ Eine Variable wird also erzeugt, indem zunächst die Objekte spezifiziert werden, auf die sich die Variable beziehen soll (siehe Abb. 3.2). Dies können Individuen, Firmen, Staaten usw. sein, wobei jedoch in den Sozialwissenschaften das Individuum in der Regel das Objekt ist. In Längsschnittdaten können auch Episoden, Übergänge oder Zeiträume die Objekte sein.

Danach wird der Merkmalsraum festgelegt. Dies bedeutet, dass „eine Menge von sich gegenseitig ausschließenden Merkmalen [bestimmt wird], durch die die Objekte, auf die man sich beziehen möchte, charakterisiert werden können“ (Rohwer & Pötter, 2002). Bei der Erhebung von Ereignisdaten können sich die Objekte in einem von (mindestens zwei) festgelegten, gegenseitig ausschließenden Zuständen befinden, die einen Zustandsraum bilden, zum Beispiel „ledig – verheiratet – geschieden – verwitwet“.

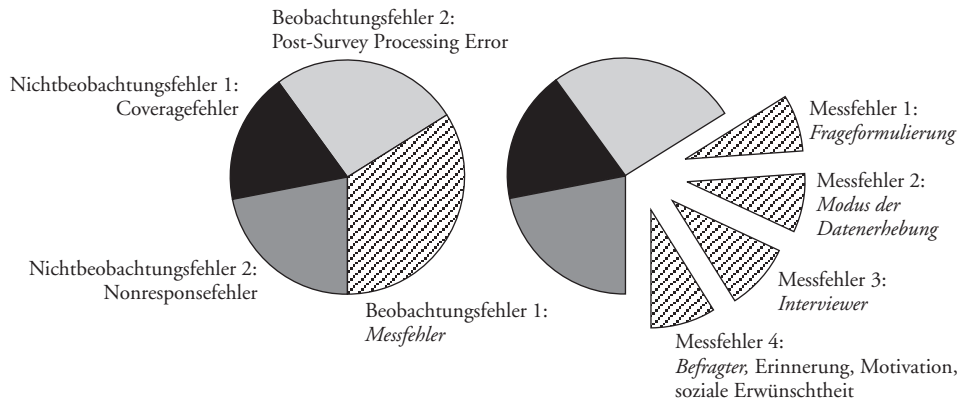
Sind das Objekt und die Merkmals- bzw. Zustandsräume festgelegt, folgt die Ermittlung der Objekte (Stichprobenziehung) und zuletzt die Zuweisung der Werte der Variablen für dieselben (Messung), beispielsweise durch eine Befragung. In standardisierten sozialwissenschaftlichen Befragungen entsprechen die Fragen meist den festgelegten Merkmals- bzw. Zustandsräumen, das heißt, die Personen werden direkt nach den Werten der statistischen Variablen gefragt.

Die Mehrheit der methodischen Literatur konzentriert sich auf die Fehler, die durch Stichprobenziehung und Messung entstehen, während die Objekt- und Variablendefinition als ein durch die Natur der zu betrachtenden Realitätsausschnitte gesetzter und in der Datenerhebung einzulösender Anspruch betrachtet wird. Allerdings sprechen sich Rohwer und Pötter (2002) für eine Terminologie und Methodologie aus, die den interessen- und vorannahmungeleiteten Charakter der Datenkonstruktion betonen. Denn am Beginn der Datenkonstruktion stehen die Annahmen der Wissenschaftler über die gesellschaftlichen Verhältnisse, die in der Objektwahl und Variablendefinition umgesetzt werden und dadurch die Anforderungen an die Rekonstruktionsarbeit der Befragten festlegen. Wenn nun die gewählten Realitätsausschnitte für die befragte Population oder generell nicht sinnvoll oder nicht angemessen sind, können auch hier Fehler entstehen.

### 3.2.2 Der Gesamtbefragungsfehler und seine Komponenten

Welche Faktoren müssen in einem Mess- bzw. Datenkonstruktionsmodell für die retrospektive Erhebung von Ereignissequenzen in Befragungen aufgenommen werden, und wie ordnet sich das Gedächtnis darunter ein? In seiner Zusammenschau der theoretischen und empirischen Fehlerliteratur aus Ökonometrie, Psychologie und Befragungsmethodologie unter-

Abbildung 3.3: Die Komponenten des Gesamtbefragungsfehlers



Der Gesamtbefragungsfehler setzt sich zusammen aus Nichtbeobachtungs- und Beobachtungsfehler. Gedächtnisfehler gehören zu den Beobachtungsfehlern.

Quelle: Groves, 1989.

scheidet Groves (1989) zwei große Gruppen von Fehlern, die im Datenkonstruktionsprozess entstehen können und die zusammen den Gesamtbefragungsfehler (*total survey error*) ergeben: den *Beobachtungsfehler* und den *Nichtbeobachtungsfehler* (siehe Abb. 3.3). Nichtbeobachtungsfehler entsprechen in der Unterscheidung von Rohwer und Pötter (2002) dem Schritt 3(a) in Abbildung 3.2, in dem die Objekte identifiziert und kontaktiert werden, also der Stichprobendefinition und -ziehung. Der Beobachtungsfehler hingegen entspricht dem Schritt 3(b), in dem die Ermittlung der Werte der Variablen für die ermittelten Objekte getätigt wird.

Groves unterteilt die Beobachtungsfehler wiederum in Messfehler und den „Post-Survey Processing Error“. Letzterer wird, wie sein Name schon sagt, durch die Weiterverarbeitung der Daten nach der Erhebung eingeführt, etwa durch Tippfehler bei der Eingabe in die Datenbanken, oder wenn fehlende Werte unangemessen imputiert werden. Der Messfehler hingegen kommt in der Befragungssituation selbst im Verlauf des Werteermittlungsprozesses zu Stande. In der Literatur werden vor allem vier Aspekte des Werteermittlungsprozesses behandelt, die das Verhältnis von tatsächlichem Sachverhalt und Datumswert beeinflussen können: der *Interviewer*, das *Erhebungsinstrument*, das *Befragungsmedium* und der *Befragte* selbst. Und hier findet sich das Gedächtnis als eine der mit dem Befragten assoziierten Fehlerquellen.

Jede der Fehlerquellen führt, wenn sie wirksam wird, erstens zu einer Zunahme des Gesamtumfragefehlers und zweitens zu Fehlern charakteristischer Art im Hinblick auf Verbreitung, Größe, Richtung und Abhängigkeit. Dies gilt für jede Befragung, ganz gleich, ob die Fragen sich auf Gegenwart oder Vergangenheit beziehen – mit Ausnahme der Erinnerungsfehlerkomponente, die nur retrospektiv erhobene Informationen betrifft. Diese kann also in retrospektiven Befragungen erstens zu einer Zunahme an Fehlern gegenüber nichtretrospektiven Befragungen führen, und zweitens zu Fehlern von charakteristischer, „gedächtnisför-

miger“ Art, die in bestimmte Richtungen weisen oder bestimmte Inhalte oder bestimmte Personen auf ganz bestimmte Weise betreffen<sup>4</sup>.

Fehler in Befragungsdaten entstehen also durch eine Vielzahl von Faktoren, die nur teilweise mit dem Antwortverhalten des Befragten zu tun haben. Und auch unter den Faktoren, die das Antwortverhalten des Befragten beeinflussen, ist die Erinnerung nur einer von mehreren. Um Abweichungen als Folge von Erinnerungsprozessen zu interpretieren, müssen also stets differenzierte Annahmen darüber entwickelt werden, welche Abweichungen tatsächlich als Resultat von Gedächtnisprozessen zu erwarten sind.

### 3.2.3 Klassische und probabilistische Testtheorie

Ausgearbeitete Modelle für die Messfehlerkomponente in Selbstauskünften wurden in der Psychologie für die Konstruktion und Interpretation einer besonderen Erhebungsform, den psychometrischen Test, entwickelt und unter den Namen klassische Testtheorie (Lord & Novick, 1968) und probabilistische Testtheorie (Krantz, Suppes, & Tversky, 1971) bekannt. Sie zeichnen sich dadurch aus, dass sie Fehler in Erhebungen explizit thematisieren und als unvermeidbar betrachten. Die Messmodelle der klassischen und der probabilistischen Testtheorie enthalten Annahmen über den Datengenerierungsprozess und das Verhältnis vom tatsächlichen zum in den Daten repräsentierten Sachverhalt. Darüber hinaus wurden aus den Modellvorstellungen Anweisungen abgeleitet, wie Tests konstruiert werden müssen, um den Voraussetzungen für das Zutreffen des jeweiligen Modells zu genügen und den Fehleranteil in den jeweiligen Daten gering zu halten. Und zuletzt beinhalten beide operationale Definitionen des Fehlers, das heißt, sie liefern Berechnungsanweisungen und Bewertungsstandards, nach denen die Fehlerhaftigkeit einer Datenmenge beschrieben und bewertet werden soll.

In der klassischen Testtheorie wird jeder erhobene Datenwert prinzipiell von zwei Komponenten als verursacht angesehen, nämlich dem realen Sachverhalt und einer Fehlerkomponente, die voneinander unabhängig auftreten und sich additiv zueinander verhalten. In der probabilistischen Testtheorie ist das Verhältnis von tatsächlichem Sachverhalt zum Datenwert ein stochastisches, das heißt, für eine Frage bzw. ein Testitem kann und muss die Wahrscheinlichkeit angegeben werden, mit der ein bestimmter tatsächlicher Sachverhalt zu einem bestimmten Datenwert führt. In beiden Fällen wird angenommen, dass diese Fehlerkomponente Fehler verursacht, die nicht bevorzugt in die eine oder andere Richtung auftreten und für alle Ausprägungen des tatsächlichen Werts gleich ausfallen.

---

<sup>4</sup> Möglicherweise beeinflussen auch andere Komponenten des Messfehlers die Erhebung von kürzer und länger zurückliegenden Sachverhalten in unterschiedlichem Maße. Retrospektive Angaben könnten dabei sogar besser abschneiden. Beispielsweise wurde verschiedentlich nachgewiesen, dass Befragte wenig vorteilhafte Ereignisse (betrunken Auto fahren, nicht zur Wahl gehen, aber auch zu hohe und zu niedrige Einkommen) bewusst verschweigen oder vorteilhafter darstellen (Belli, Shay, & Stafford, 1999; Johnson, Gerstein, & Rasinski, 1998). Diese Effekte der sozialen Erwünschtheit könnten für retrospektive Angaben *schwächer* ausfallen, da es den Befragten vermutlich weniger wichtig ist, sich auch in der Vergangenheit in einem guten Lichte zu präsentieren. Untersuchungen hierzu liegen aber meines Wissens nicht vor. Unter den Gedächtnisfehler kann dies allerdings nicht gerechnet werden, da es sich um ein absichtsvolles Verschweigen oder Beschönigen von Information handelt.

Allerdings kritisieren Rohwer und Pötter (2002) die Übertragung testtheoretischer Annahmen und Termini auf sozialwissenschaftliche Befragungsdaten. Die Bezeichnung des Erhebungsprozesses als „Messung“ erfolge unter expliziter oder impliziter Bezugnahme auf ein repräsentationales Messkonzept, das aus der Physik über die Psychophysik und die Psychometrie auch in den Sozialwissenschaften Einzug gefunden habe. Damit einher gehe eine irreführende Bezugnahme auf die Fehlertheorie und -terminologie, die in diesem Rahmen entwickelt wurde, namentlich eine Übernahme von Reliabilität und Validität in ihren inhaltlichen und operationalen Definitionen (siehe Abschnitt 3.3).

Zwar sähen sich Tests und Befragungen im äußeren Erscheinungsbild oft sehr ähnlich, da sie beide eine Reihe von Fragen zu Ansichten oder Verhaltensweisen enthalten. Rohwer und Pötter stellen aber mit Nachdruck heraus, dass Tests und Befragungen auf grundlegend unterschiedlichen Annahmen über den Datenkonstruktionsprozess beruhen. Psychometrische Tests sollen ein latentes Konstrukt (beispielsweise Intelligenz, Depression oder Lebenszufriedenheit) erfassen, über das die betroffenen Personen nicht direkt Auskunft geben können. Testitems werden dabei als Stimuli betrachtet, die beim Testanwender behaviorale Manifestationen dieses Konstrukts, nämlich seine Antworten, hervorrufen. Personen mit unterschiedlichen Ausprägungen des latenten Konstrukts reagieren dabei unterschiedlich, sodass unterschiedliche Antworten auf Ausprägungsunterschiede im zu Grunde liegenden Konstrukt hindeuten und also über die Ausprägung desselben beim Testanwender Auskunft geben. Deshalb sind zur Messung eines einzigen latenten Konstrukts mehrere Testitems nötig, damit sich der Messfehler (der ja zufällig und unabhängig vom wahren Wert auftritt) für jede einzelne Person über alle Items „wegmitteln“ und die getestete Person anhand der Gesamtheit aller behavioralen Manifestationen hinsichtlich ihrer Ausprägung auf dem latenten Konstrukt eingeschätzt werden kann.

Befragungen werden hingegen als ein Mittel betrachtet, den Befragten zur Abgabe von Information zu veranlassen, über die nur er alleine verfügt und auch Auskunft geben kann. Befragungsfragen stellen dem Befragten in sprachlicher Form die Information zur Verfügung, die er benötigt, die Anforderungen seines „Gesprächspartners“ zu verstehen und so gut wie möglich zu erfüllen. Jeder Sachverhalt wird dabei durch nur eine Frage erfasst, die direkt auf den interessierenden Realitätsausschnitt zielt. Befragungsdaten werden also als Ergebnisse eines Prozesses sozialer Informationsverarbeitung von Menschen konzipiert, das heißt dass kognitive und interpersonelle Prozesse an der Datenproduktion beteiligt sind (Schwarz & Sudman, 1996; Sudman u.a., 1996; Tourangeau, Rips, & Rasinski, 2000)<sup>5</sup>. Deshalb können die Annahmen und Datengenerierungsmodelle der Psychometrie und besonders die Definitionen von Reliabilität und Validität nicht auf Befragungsdaten übertragen werden.

Im folgenden Abschnitt möchte ich einige Ansätze vorstellen, die sich mit Befragungen als Situation sozialer Informationsverarbeitung befassen und diese kognitiven und interpersonellen Prozesse beschreiben.

---

<sup>5</sup> Ob dieses Grundmodell nicht doch auch auf psychometrische Tests in stärkerem Umfang zutrifft als mit den Modellvorstellungen der klassischen und probabilistischen Testtheorie vereinbar, wird verschiedentlich diskutiert (Hamilton & Shuminsky, 1990; Knowles, 1988).

### 3.2.4 Kognitive Konzepte der Befragung

Sudman und Bradburn (1974) beschreiben die Erhebungssituation als eine Sequenz kognitiver Aufgaben, die vom Befragten ausgeführt werden sollen. Diese Aufgaben werden in den Instruktionen und Fragen an ihn herangetragen. Diese enthalten die Information darüber, was der Befragte tun soll und wie er es tun soll. Die Befragten sind dabei aktiv an der Datenproduktion beteiligt und nutzen die in den Fragen enthaltene Information, um die von ihnen geforderte kognitive Aufgabe zu verstehen und zu bewältigen. Abbildung 3.4 zeigt ein Modell nach Lessler und Forsyth (1996) und Sudman u.a. (1996), das für die Beantwortung von Einstellungs- und Meinungsfragen mit Ratingskalen oder Antwortkategorien entwickelt wurde. Der Befragte versucht hier zunächst, die an ihn gestellte Anforderung zu verstehen und eine plausible Interpretation dessen zu konstruieren, was sein Gegenüber von ihm erfahren möchte (Schritt 1). Danach ruft er, ausgerüstet mit dieser Interpretation, die erforderliche Information ab bzw. rekonstruiert sie (Schritt 2), wobei er seine Rekonstruktionen zwischen- und endevaluiert (Schritt 3). Dann „editiert“ er seine Antwort indem er überprüft, ob er die rekonstruierte Information nicht aus irgendwelchen Gründen doch verschweigen möchte (Schritt 4). Zuletzt bringt er sie in die vom Wissenschaftler vorgegebene Form – etwa einer numerischen oder kategorialen Ratingskala, wenn etwa angegeben werden soll, ob man eine bestimmte Verhaltensweise „oft – manchmal – selten – nie“ ausführt oder einer Aussage „sehr zustimmt – eher zustimmt – eher ablehnt – sehr ablehnt“ (Schritt 5). Bei jedem dieser Schritte können Abweichungen der Angaben von der Realität entstehen.

Für Gedächtnisfehler in Ereignisdaten ist das Modell nicht direkt übertragbar. Fehler bei Schritt 4 (Editierung) fallen nicht unter die Gedächtnisfehlerdefinition, da es sich hierbei um absichtsvolles Verschweigen handelt. Schritt 5 (Formatierung) muss nur ausgeführt werden, wenn Ratingskalen eingesetzt werden. Die Schritte 1 bis 3 sind allerdings auch bei der retrospektiven Ereignisdatenerhebung relevant, wobei die Unterscheidung zwischen Erinnern und Evaluierung nicht getroffen werden muss. Es wäre also zu fragen, ob und wie bei der retrospektiven Rekonstruktion Fehler dadurch zu- oder eine andere Form annehmen, dass Fragen anders interpretiert werden und Inhalte anders (oder nicht mehr) rekonstruiert werden können.

Die Frage, ob Inhalte zu späteren Zeitpunkten anders oder nicht mehr erinnert werden können, habe ich ausführlich in Kapitel 2 behandelt. Was die Interpretation der Frage anbetrifft, so haben kognitiv begründete Methoden zur Überprüfung des Frageverständnisses belegt, dass auch scheinbar unproblematische Begriffe von Befragten sehr unterschiedlich interpretiert werden (Fowler & Roman, 1992; Groves, 1996; Sudman u.a., 1996) – etwa

Abbildung 3.4: Datengenerierungsmodell für Einstellungsurteile in Befragungen

---

Schritt 1	Verstehen bzw. Interpretieren der Frage
Schritt 2	Erinnern/Abrufen von relevanten Inhalten
Schritt 3	Evaluation der gefundenen Inhalte
Schritt 4	Editierung
Schritt 5	Formatierung

---

Quelle: Lessler & Forsyth, 1996; Sudman, Bradburn, & Schwarz, 1996.

„im letzten Jahr“ oder „Arztbesuch“, ganz zu schweigen von „Haushaltsnettoeinkommen“ (Kühnemund, 1990). Es kommt vor, dass die Befragten sogar selbst definieren sollen, ob ein bestimmter Zustand zum Beispiel als Arbeitslosigkeit oder feste Partnerschaft gilt. Oder es gibt prinzipiell eindeutige und unmissverständliche Definitionen für ihre Zustände, die aber zu kompliziert sind, als dass sie dem Befragten in vollem Umfang mitgeteilt werden könnten. In solchen Fällen besteht ein erhöhtes Risiko, dass die Interpretation von dem abweicht, was der Wissenschaftler im Sinn hatte, und dass sie zu zwei Zeitpunkten abweichend ausfällt. Auch wenn Definitionen gegeben werden, kann man sich nicht darauf verlassen, dass eine Frage wirklich immer im Sinne des Wissenschaftlers interpretiert wird, besonders wenn die Interpretation komplex ist und dem Alltagsverständnis des Befragten zuwiderläuft. Wenn in retrospektiven Befragungen die zur Selbstbeschreibung angebotenen Zustände zu späteren Zeitpunkten oder für weiter zurückliegende Zeiträume anders angewendet werden, können abweichende Interpretationen ebenfalls als Folge der verstrichenen Zeit auftreten und als Gedächtnisfehler bezeichnet werden.

Zusammenfassend kann man sagen, dass die Beschreibung und Erklärung von gedächtnisverursachten Fehlern in retrospektiv erhobenen Ereignisdaten eine komplexe und multidimensionale Aufgabe darstellt. Zahlreiche miteinander zusammenhängende Faktoren können für Abweichungen von Befragungsdaten von der Realität verantwortlich sein und deren Verbreitung, Ausmaß, Richtung und Abhängigkeit beeinflussen. In kontrollierten Studien wurde der Einfluss einzelner Fehlerkomponenten getrennt untersucht, indem der Einfluss der anderen konstant gehalten oder auf andere Weise kontrolliert wurde; einen Überblick gibt Groves (1989). Bei bereits erhobenen Befragungsdaten lässt sich dies allerdings nicht leisten. Im nächsten Abschnitt möchte ich Konzepte und Methoden darstellen, mit denen der Fehler dennoch eingeschätzt werden kann und herausarbeiten, unter welchen Bedingungen die so gefundenen Abweichungen im Hinblick auf die Gedächtnisfehlerkomponente interpretiert werden können.

### 3.3 Ermittlung und Quantifizierung des Gedächtnisfehlers in Befragungsdaten

#### 3.3.1 Reliabilität und Validität in Befragungen

##### *Konzept und Ermittlung*

Verbreitet ist der Wunsch, die Fehlerhaftigkeit einer konkreten Datenmenge in einem quantitativen Gütekennwert zu charakterisieren, um deren Aussagekraft im Vergleich mit anderen Datenmengen oder einem Standard zu bewerten. Als zentrale Gütekennwerte für Daten gelten *Validität* und *Reliabilität*. Dabei handelt es sich um theoretische Konstrukte, anhand derer sich Datenmengen charakterisieren lassen: Reliabilität, auch Zuverlässigkeit genannt, bezeichnet das Ausmaß, in dem ein an einem Objekt mit einer Erhebungs- oder Messprozedur erhobener Datumswert bei einer erneuten Anwendung derselben Erhebungsprozedur am selben Objekt wieder denselben Datumswert ergibt. Dies verweist auf die Eigenschaft der Erhebungsprozedur, ein und denselben tatsächlichen Sachverhalt auch durch ein und denselben Datenwert zu repräsentieren, und somit auf die Güte der mit dieser Methode generierten Daten. Validität bezeichnet den Grad, in dem eine Erhebungsprozedur tatsächlich das Konzept erfasst, das der Wissenschaftler erheben möchte, das heißt wie gut die Daten mit der Realität übereinstimmen (Bortz, 1993; Lord & Novick, 1968). Die Reliabilität ist eine notwendige, nicht aber eine hinreichende

Bedingung für Validität, da Daten, die bei verschiedenen Gelegenheiten valide erhoben wurden, notwendigerweise auch reliabel, das heißt übereinstimmend über diese Gelegenheiten hinweg sind. Umgekehrt sind reliable Daten nicht notwendigerweise auch valide: Der reliable, wiederholt identisch erhobene Wert kann auch wiederholt falsch erhoben worden sein. Reliabilität und Validität sind dabei immer Eigenschaften einer Datenmenge, nicht eines einzelnen Datums – eine individuelle Antwort kann also weder reliabel noch valide sein (Rohwer & Pötter, 2002)<sup>6</sup>.

Rohwer und Pötter (2002) kritisieren diese Konzepte von Reliabilität und Validität als unsinnig und irreführend. Ihre Kritik zur Reliabilität bezieht sich vor allem auf die vorausgesetzte, aber gerade in Längsschnittdaten nicht gegebene Merkmalskonstanz – wobei es auch Reliabilitäts- und Übereinstimmungskonzepte für zyklisch veränderliche Merkmale und Sequenzen gibt (Abbott & Tsay, 2000; Dijkstra & Taris, 1995). Zur Validität meinen die Autoren, dass es „nicht darum [geht], ob ein Messinstrument ‚das misst, was es messen soll‘ (diese Frage könnte auch nicht beantwortet werden), sondern es geht darum, ob bzw. wie gut sich ein Messinstrument eignet, bestimmte Zwecke zu erreichen“ (Rohwer & Pötter, 2002, S. 124). Da die Praxis der Validitätsprüfung in vollem Umfang auf eine Zweckmäßigkeitprüfung hinauslaufe (etwa die Brauchbarkeit von Daten, andere Daten vorherzusagen), vertreten die Autoren die Auffassung, dass der Zweckmäßigkeitgedanke auch auf der konzeptuellen Ebene an die Stelle des Validitätsgedankens treten und zur Bewertungsgrundlage werden solle.

Die *Reliabilität* von Befragungsdaten wird durch den Vergleich von mehrmals durchgeführten Erhebungen derselben Information an denselben Personen in vergleichbaren Befragungen ermittelt<sup>7</sup>. Zur *Validitätsschätzung* wird ein Vergleich der zu validierenden Befragungsdaten mit anderen, unabhängig davon erhobenen Daten durchgeführt. Diese anderen Daten gelten als weniger fehlerbehaftet und werden als Kriterium verwendet, an dem die Befragungsdaten gemessen werden können. Die Maßzahlen für Reliabilität und Validität einer Datenmenge sind dann Maße der Übereinstimmung bzw. Abweichung.

Die Aussagekraft einer einzigen Kennzahl ist allerdings begrenzt; Kruskal (1991) meint sogar: „The idea of a single number reflecting error magnitude of an observational system has a powerful appeal, especially if we are comparing two or more observational systems. Yet the use of a single figure of merit can be misleading.“ Erstens ist am Fehler mehr als nur ein Aspekt von Interesse: Wie in Abschnitt 3.1.2 dargestellt, lassen sich die Fehler in einer Datenmenge hinsichtlich Verbreitung, Ausmaß, Richtung und Abhängigkeit vom tatsächlichen Sachverhalt

<sup>6</sup> Häufig bleibt in der Terminologie zweideutig, ob Reliabilität und Validität Charakteristika einer Datenmenge oder einer Erhebungsprozedur sind (Sawilowski, 2000; Thompson, 1994; Thompson & Vacha-Haase, 2000). Tatsächlich jedoch sind die genannten Größen immer Kennzeichen einer und nur einer Datenmenge. Ein Reliabilitäts- oder Validitätswert kann über die Erhebungsprozedur höchstens sagen, dass sie für die konkrete Stichprobe reliable bzw. valide Daten produzieren kann; von der Reliabilität eines Erhebungsverfahrens zu sprechen wäre irreführend (Thompson & Vacha-Haase, 2000). Mit der üblichen methodischen und interpretatorischen Vorsicht kann aber eine Generalisierung von einer Datenmenge auf andere unter vergleichbaren Bedingungen an vergleichbaren Personen gewonnene Datenmengen getätigt werden. Ob diese Vergleichbarkeit gegeben ist, kann nur entschieden werden, wenn die Datengenerierungsprozesse und ihre fehlerrelevanten Eigenschaften bekannt sind.

<sup>7</sup> Für andere Datenerhebungsprozeduren als Befragungen gibt es weitere Möglichkeiten, Reliabilität zu schätzen. In psychometrischen Tests, wo ein und dieselbe latente Größe durch mehrere Items erfasst wird, kann man beispielsweise die eine Hälfte dieser Items mit den anderen vergleichen (die *Split-Half*- oder Testhalbierungs-Reliabilität). In Beobachtungsstudien werden Übereinstimmungsmaßzahlen zwischen zwei oder mehreren Beobachtern berechnet und an der zufällig zu erwartenden Übereinstimmung relativiert.



charakterisieren. Dies würde bereits vier Maßzahlen je Informationsaspekt erforderlich machen.

Zweitens können für verschiedene Gruppen von Personen unterschiedliche Reliabilitäten und Validitäten vorliegen, wenn zum Beispiel Frauen bessere Angaben machen als Männer oder Jüngere bessere als Ältere. Praktisch kann dies auch darauf hindeuten, dass die festgelegten Merkmals- oder Zustandsräume für bestimmte Personen weniger angemessen sind als für andere. Drittens gibt es viele Informationsaspekte in einem Datensatz, deren Übereinstimmung sehr unterschiedlich ausfallen kann. Durch Aggregation, Rangskalierung oder Zusammenlegung von Kategorien können beliebig viele weitere Vergleichsinformationen gebildet werden. Durch eine einzige Kennzahl kann man jedoch nicht allen gerecht werden.

Auf der Grundlage einer einzigen oder einiger weniger Maßzahlen kann also kein differenziertes und abgewogenes Urteil darüber gefällt werden, für welche Analysen die Daten uneingeschränkt geeignet sind und für welche Personen, Informationen und Analysen welche Einschränkungen getroffen werden müssen. Studien zur Abschätzung von Datenfehlern in Befragungsdaten berichten in der Regel eine Reihe von Übereinstimmungsinformationen. Anders als beispielsweise in der klassischen Testtheorie gibt es aber keine verbindlichen Regelungen über die Vergleichsprozedur, Maßzahl(en) und deren Bewertung<sup>8</sup>; geprüft werden meist die Informationsdimensionen, die für die inhaltlichen Interessen der Studiendurchführer am relevantesten sind und für die Schwierigkeiten angenommen werden. In den nächsten zwei Abschnitten werde ich einen Überblick über die verwendeten Übereinstimmungsmaße geben.

#### *Vergleiche auf Individualebene*

Auf der Individualebene werden meist einfache Übereinstimmungsmaße berichtet, zum Beispiel der Anteil der übereinstimmenden bzw. abweichenden Angaben an allen Angabenpaaren. Dieses Maß deckt vor allem die Verbreitung des Fehlers ab, sagt aber nichts über Größe, Richtung oder Abhängigkeit der Fehler und differenziert nicht nach Irrtumsrisiken und Fehldiagnoserisiken. Klein und Fischer-Kerli (2000) betrachten die Reliabilität von Angaben zur Partnerschaftsgeschichte in Analogie zu einer Beurteilerübereinstimmung und relativieren die beobachteten Übereinstimmungsraten an dem Ausmaß zufällig erwarteter Übereinstimmung. Hier erlaubt die Charakterisierung von Fehlern als *hits* und *misses* eine Unterscheidung von Irrtümern und Fehldiagnosen.

Für metrische Variablen wie Anzahl und Dauer von Ereignissen oder auch Einkommensangaben wird oft die Korrelation der beiden Angaben berechnet oder die mittlere numerische Abweichung aller Angabenpaare mit oder ohne Vorzeichen angegeben, teilweise unter Angabe von Signifikanzen (Belli, Shay, & Stefford, 2001; Mathiowetz & Duncan, 1988; Paull, 2002). Eine hohe Korrelation besteht beispielsweise dann, wenn diejenigen, die innerhalb der Stichprobe beim ersten Mal die geringeren Einkommen hatten, auch in der zweiten Befragung inner-

<sup>8</sup> In der klassischen Testtheorie wird Reliabilität operational definiert als die Korrelation der zu zwei oder mehr Erhebungszeitpunkten erhobenen Messungen an einer Stichprobe und Validität als die Korrelation mit dem tatsächlichen Wert. Auch Bewertungsstandards werden vorgegeben: Eine gute Messung sollte Reliabilitätskoeffizienten von mindestens .8 aufweisen, wobei Reliabilitäten zwischen .8 und .9 als mittelmäßig, Reliabilitäten über .9 als hoch gelten (Bortz & Döring, 1995). Da sich alle psychometrischen Tests in der Ermittlung ihrer Reliabilitätskoeffizienten an diese operationalen Vorgaben halten, werden diese vergleichbar.

halb des Einkommensspektrums zu den Niedrigverdienern gehören, während zugleich die Besserverdienenden aus der ersten Befragung auch in der zweiten Befragung höhere Einkommen angeben – auch wenn die einzelnen Einkommensangaben voneinander abweichen (Verbreitung des Fehlers) und Abweichungen sehr groß ausfallen (Fehlergröße)<sup>9</sup>. Korrelationen informieren also hauptsächlich darüber, ob die Abweichungen alle in dieselbe Richtung auftreten.

Was sagt es uns über den Fehler, wenn Differenzen mit einem t-Test für paarweise Vergleiche auf Signifikanz überprüft werden? Die numerischen Differenzen zwischen den Angaben können positiv ausfallen (wenn der Ausgangsbetrag größer ist als der davon abgezogene) oder negativ ausfallen (wenn der subtrahierte Betrag größer ist als der Ausgangsbetrag). Der t-Test prüft die Wahrscheinlichkeit, dass die Differenzen zwischen den Angabenpaaren im Mittel verschieden von Null sind. Außer in dem unwahrscheinlichen Fall, in dem alle Angaben perfekt übereinstimmen, ist dies der Fall, wenn die Differenzen im selben Maße positiv und negativ ausfallen und sich somit im Mittel zu Null addieren. Damit wird der t-Test zu einer Überprüfung der Fehlerrichtung: Ein signifikantes Ergebnis bedeutet, dass die Abweichungen sich nicht gegeneinander aufheben.

Wenn die numerischen Differenzen ohne Vorzeichen betrachtet werden, ist dies ein Maß für die Größe der Abweichungen, ein t-Test würde hier allerdings weniger Sinn machen, da ja nie die Annahme bestand, die mittlere vorzeichenfreie Abweichung könnte Null betragen.

Für den Vergleich von Angaben auf der Individualebene ist die Zuordnung der Informationen aus den beiden zu vergleichenden Datensätzen (also der Befragtenangaben zu den externen Daten bzw. der beiden Befragtenangaben aus Wiederholungsbefragungen zueinander) für jedes Individuum bzw. jedes Objekt erforderlich. Dies ist ein komplexer Vorgang und kann nicht immer in befriedigender Weise geleistet werden (mehr hierzu in Abschnitt 3.3.5). Deshalb werden Vergleiche auch häufig auf Stichprobenebene angestellt.

#### *Vergleiche auf Stichprobenebene*

Andere Vergleiche ermitteln Reliabilität und Validität, indem die beiden zu vergleichenden Datenmengen getrennt betrachtet und Stichprobenkennwerte über alle Teilnehmer hinweg miteinander verglichen werden, beispielsweise Mittelwerte, Mediane oder Verteilungen. Verschiedentlich werden auch Koeffizienten von Regressionsmodellen verglichen, die mit den Daten der zu vergleichenden Datensätze getrennt geschätzt wurden. Hier lässt sich ein Datensatz als reliabel bzw. valide bezeichnen, wenn die Modellkoeffizienten in beiden Datensätzen übereinstimmend signifikant bzw. nichtsignifikant ausfallen und in dieselbe Richtung zeigen oder allgemeiner, wenn die Schlussfolgerungen sich nicht in bedenklicher Weise voneinander unterscheiden. Es deutet schon auf ein Konzept von Datengüte im Sinne von Zweckmäßigkeit oder Brauchbarkeit hin, wenn Schömann (1994) schreibt: „We want to recast the reliability question by defining retrospective data to be sufficiently reliable if (...) coefficients and their level of significance lead to the same conclusions in both analyses.“ (S. 159)

<sup>9</sup> Von einem pragmatischen Standpunkt sagt dieses Maß allerdings auch direkt aus, dass anhand der Daten die Niedrig- von den Besserverdienenden zuverlässig und in jeder Befragung übereinstimmend unterschieden werden können. So kann ein eingeschränkt aussagekräftiges Reliabilitätsmaß gleichzeitig ein Maß für die Zweckmäßigkeit der Daten für bestimmte Interessen darstellen, wie sie von Rohwer und Pötter (2002) auch im Rahmen von Reliabilitäts- und Validitätsschätzungen befürwortet werden.

Angaben über die Reliabilität bzw. Validität von Aggregateigenschaften und Modellkoeffizienten sind für den Sozialwissenschaftler von unmittelbarerem Interesse als eine genaue Auskunft über Abweichung auf der Ebene der einzelnen Erhebung. Allerdings besteht die Gefahr, dass die Übereinstimmung im Aggregatvergleich oder beim Vergleich von Modellkoeffizienten überschätzt wird, weil Fehler der Individuen sich gegeneinander aufheben können. Wenn die Hälfte der Teilnehmer in einer späteren Befragung einen zu hohen, die andere Hälfte hingegen einen zu niedrigen Wert angibt, kann der Mittelwert der Stichproben perfekt übereinstimmen. Daraus aber auf die Fehlerfreiheit der Angaben zu schließen oder anzunehmen, dass auch alle anderen Analysen mit den Daten genauso unberührt vom Gedächtnisfehler sind, wäre falsch. Tatsächlich kommen Vergleiche auf Stichprobenebene in der Regel zu positiveren Urteilen über die Datenqualität als solche, die die Daten auf der Individualebene vergleichen (Jacobs, 2002). Vergleiche auf der Individualebene sind also härtere Prüfungen der Datengüte. Aggregatvergleiche oder getrennte Modellschätzungen können ergänzend durchgeführt werden, um zusätzlich zum Fehler auch seine Auswirkung abschätzen zu können, um damit eine Art „pragmatische Datengüte“ oder Zweckmäßigkeit zu erfassen.

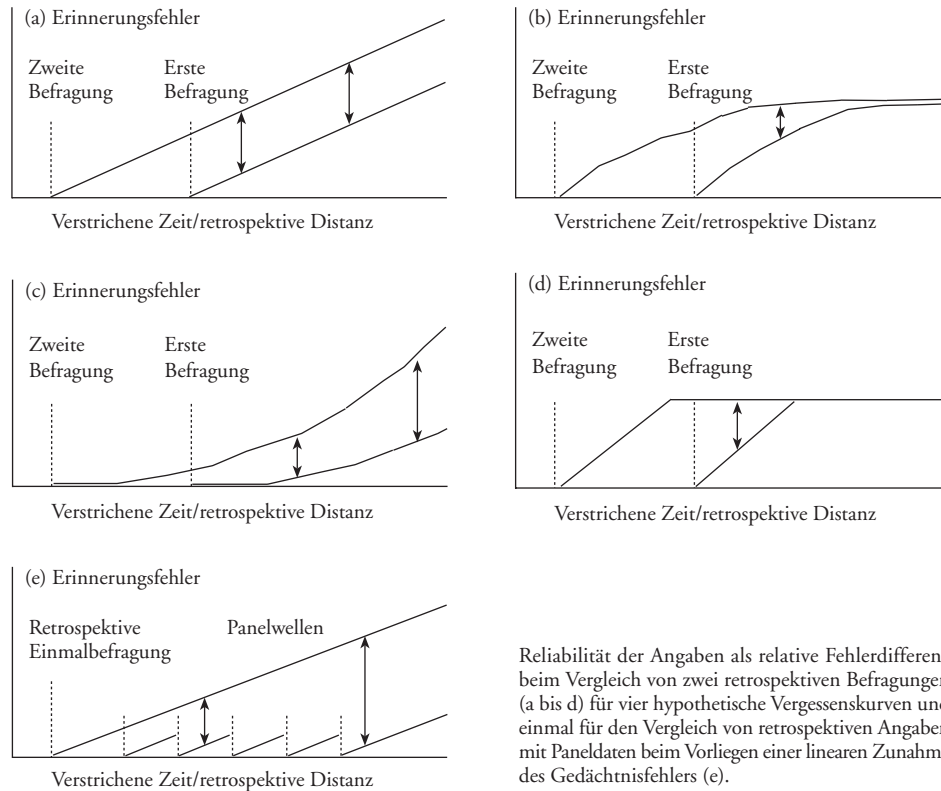
Dieser Abschnitt hat gezeigt, dass es viele Möglichkeiten gibt, Reliabilität und Validität in Befragungen zu quantifizieren, die in sehr unterschiedlichem Maße über den Fehler und seine verschiedenen Aspekte Auskunft geben. Keinesfalls gibt es ein Standardvorgehen, wie es durch die operationale Definition von Reliabilität und Validität aus der klassischen Testtheorie abgeleitet wurde. Im nächsten Absatz werde ich untersuchen, wie sich verschiedene Methoden der Reliabilitäts- und Validitätsschätzungen als Gedächtnisfehlerprüfungen einsetzen lassen.

### 3.3.2 Reliabilitätsstudien: Die Wiederholungsbefragung

Wenn zur Abschätzung der Reliabilität Wiederholungsbefragungen durchgeführt werden, muss dabei zwischen den Befragungen ein angemessen langer Zeitraum liegen. Bei zu kurzen Zeiträumen könnten sich die Befragten nicht an den abgefragten Sachverhalt, sondern an ihre Angaben aus der letzten Befragung erinnern und diese wiederholen – oder sogar absichtlich etwas anderes angeben, wenn sie annehmen, dass der Wissenschaftler nun auf etwas anderes hinauswill (Sudman u.a., 1996). Wir setzen voraus, dass der Erinnerungsfehler bei Statusangaben über die Gegenwart Null ist und mit zunehmender retrospektiver Distanz zunimmt. Insofern werden Angaben, die zu einem späteren Zeitpunkt gemacht werden, fehlerhafter sein und deshalb von den (korrekteren) früheren auf charakteristische Weise abweichen. Wenn deshalb die Reliabilität einer retrospektiven Befragung durch den Vergleich von Wiederholungsbefragungen geprüft wird, sind die im vorigen Kapitel diskutierten Reliabilitätsmaße nicht nur Maße für Datenqualität. Sie kann auch als ein Maß für den durch die zusätzlich verstrichene Zeit eingeführten Fehler, das heißt den Erinnerungsfehler, interpretiert werden.

Vergleiche zwischen Wiederholungsbefragungen, die längere Zeiträume abdecken, verraten, ob die späteren Angaben fehlerhafter sind als die früheren, nicht aber, wie fehlerhaft schon die früheren Angaben durch die dort gemachten Erinnerungsfehler waren. Ein Vergleich informiert uns also über die relative Fehlerdifferenz der beiden Angaben, und dies muss bei der Interpretation der Abweichungen berücksichtigt werden. Ausschlaggebend ist dabei, welche Beziehung zwischen sich verlängernder retrospektiver Distanz und Ausmaß des Fehlers besteht und wie sich dazu der Zwischenerhebungszeitraum verhält. In Abbildung 3.5 ist dar-

Abbildung 3.5: Gedächtnisfehlerermittlung in Wiederholungsbefragungen



gestellt, wie diese relative Fehlerdifferenz für längere Zeiträume zu interpretieren ist. Die senkrechten Pfeile stehen für das Ausmaß der relativen Fehlerdifferenz an einem bestimmten Zeitpunkt innerhalb der retrospektiven Periode. Wie groß diese ausfällt, hängt von den Kurvenparametern und der Länge des Zwischenerhebungszeitraumes ab<sup>10</sup>. Abbildung 3.5a zeigt, wie sich die Angaben zueinander verhalten, wenn die Fehlerzunahme eine lineare Funktion der verstrichenden Zeit ist, wie von Linton (1975) und Rubin und Baddeley (1989) für verschiedene Erinnerungsinhalte angenommen wird: Sie fallen über den gesamten abgedeckten Zeitraum hinweg konstant aus. Abbildung 3.5b stellt den Zusammenhang im Fall einer logarithmischen Funktion dar; hier ist die Abweichung für den Zeitpunkt des ersten Interviews am größten, nimmt dann langsam ab und ist ab irgendeinem Punkt nicht mehr nachweisbar, da die Kurven sich derselben Asymptote nähern. Auch hier bestimmen der Zwischenerhebungszeitraum und die Kurvenparameter, wie groß die relative Fehlerdifferenz zu Beginn ausfällt

<sup>10</sup> Wenn ich hier von der Zunahme des Fehlers spreche, sind damit alle vier Aspekte von Fehlerhaftigkeit gemeint: eine größere Verbreitung von Fehlern, größere Abweichungen, verstärkte Gerichtetheit oder stärkere Abhängigkeit vom tatsächlichen Sachverhalt.

und ab welchem Punkt die Kurven konvergieren. Weitere Möglichkeiten wären eine exponentielle (Abb. 3.5c) oder eine diskontinuierliche Funktion (Abb. 3.5d) wie sie von Bahrick, Bahrick und Wittlinger (1975) und Elias (1997) belegt wurden. Abbildung 3.5e zeigt den linearen Fall beim Vergleich einer retrospektiven Einmalbefragung mit Angaben aus mehreren Panelwellen, die ihre Zwischenerhebungszeiträume retrospektiv abdecken. Wenn also Angaben aus zwei retrospektiven Befragungen zu einem Zeitpunkt, der beide Male in der Vergangenheit liegt, konsistent ausfallen, kann es dennoch sein, dass beide fehlerhaft sind (vgl. Abb. 3.5b und 3.5d); wenn die Abweichungen zwischen Wiederholungsbefragungen mit der retrospektiven Distanz linear zunehmen, kann dem ein exponentiell zunehmender Erinnerungsfehler zu Grunde liegen usw.

### 3.3.3 Validierungsstudien: Die Registerdatenüberprüfung

Während Befragungen oft der beste Weg sind, von einer Stichprobe vielfältige und detaillierte Information zusammenzutragen, findet sich teilweise dieselbe Information auch in Datensätzen, die von Institutionen und Organisationen angelegt wurden, um zur Erfüllung ihrer jeweiligen Aufgaben über Individuen Buch zu führen. Beispielsweise führen Versicherungsträger über die Versicherten Buch, Steuerbehörden über Steuerzahler, Krankenhäuser über Patienten und Bildungseinrichtungen über ihre Studenten. Solche Daten werden „prozessproduzierte“ (Bick, Mann, & Müller, 1984) oder Registerdaten (Groves, 1989) genannt. Sie gelten oft als weniger fehlerhaft als Befragungsdaten, da sie nicht auf Erinnerungs- oder sonstigen kognitiven Prozessen beruhen und werden als externe Kriterien zur Validierung von Befragungsdaten herangezogen. Abweichungen zwischen Befragungs- und prozessproduzierten Daten können deshalb auch als Folge der kognitiven Prozesse in der Befragung interpretiert werden.

Die Annahme, dass Registerdaten immer besser als Befragungsdaten sind, ist allerdings nicht uneingeschränkt haltbar. Beispielsweise gelangen nicht alle angezeigten Straftaten in die Register; der aufnehmende Beamte oder die strafverfolgenden Behörden können eine Anzeige als unbegründet einstufen, woraufhin der Vorgang aus den Akten getilgt werden muss. Schwarzarbeiter werden nicht in den Personalunterlagen geführt, geben die Erwerbstätigkeit möglicherweise aber in Befragungen an. Akten können auch verloren gehen oder falsch abgelegt werden. Groves (1989) äußert außerdem die Befürchtung, dass die Beobachtungsfehler in Befragungs- und Registerdaten korreliert sind, dass also genau die Angaben, die in Befragungsdaten problematisch sind, auch in Registerdaten mangelhaft repräsentiert sind. Beispielsweise werden Episoden der Schwarzarbeit nicht in den Personalakten geführt, aber vermutlich auch in Befragungen seltener berichtet als reguläre Arbeitsverhältnisse.

Beim Vergleich von Befragungs- und Registerdaten muss also auch der Datenkonstruktionsprozess der Registerdaten bekannt sein und berücksichtigt werden, um die Abweichungen als Schwächen der Befragungsdaten interpretieren zu können. Leider sind selbst die administrativen Regel- und Vorschriftensysteme, die festlegen, welche Personen und Sachverhalte beim Vorliegen welcher Bedingungen durch welchen Prozess erfasst werden sollen, oft nicht in einer für den Sozialwissenschaftler praktisch nutzbaren Form zugänglich. Die Realitäten des Produktionsprozesses sind noch viel weniger bekannt. Bereits 1980 schreiben Bick und Müller in der Einführung zu ihrem Herausgeberband:

„Die Nutzung von prozess-produzierten Daten (...) für Forschungszwecke bedeutet Arbeit mit Material, auf das faktisch eine große Anzahl von Störfaktoren einwirkt. (...) Bislang liegen noch wenig systematische Kenntnisse über diese Einflussfaktoren vor. (...) Kennzeichnend ist vielmehr dass ein weit verstreutes und jeweils begrenztes Einzelwissen bei den verschiedenen Nutzern von prozess-produzierten Daten vorliegt (...) (dies ist) zumeist ‚Wissen in Hinterköpfen‘ und somit wie alles mündlich überlieferte Wissen ständig der Gefahr ausgesetzt, ‚auszusterben‘. Da es zumeist als Kunstfertigkeit und nicht als systematisches Wissen ausgebildet ist, sind hier die möglichen Überlieferungsverluste besonders gravierend.“

Ein konzeptueller Rahmen für Fehler in prozessproduzierten Daten findet sich bei Bick und Müller (1984). Hier verlangt die datenproduzierende Behörde ähnlich wie in einer Befragung Menschen Informationen ab, und zwar geschieht dies im Rahmen von Verwaltungskontakten und vermittelt durch Formulare und andere Aufzeichnungsvorschriften. Drei Faktorenbündel beeinflussen dabei die Güte der Datensammlung (siehe Abb. 3.6): der *Aktenproduzent* innerhalb der Institution, die „*Verwaltungstheorie*“ (d.h. die Regelungen, welche Information von welchen Personen in welcher Form in welchen Zeitabständen auf welchem Wege erhoben werden soll) und zuletzt der *Klient oder Betroffene*, der außerhalb der datensammelnden Institution steht und ihr die Information zur Anzeige bringt. Studien, die sich mit den Fehlern beschäftigen, die bei der Entstehung prozessproduzierter Daten auftreten, existieren meines Wissens kaum (eine Ausnahme stellt die Untersuchung von Bick & Müller [1980] dar).

Trotz alledem kann man davon ausgehen, dass der Fehler in prozessproduzierten Daten zumindest nicht durch die zwischen Akteneintrag und Akteineinsicht verstrichene Zeit beeinflusst wird. Durch Erinnerungsfehler in den Befragungsdaten müssten Abweichungen demnach umso größer ausfallen, je länger die Zeitpunkte, auf die die Angaben sich beziehen, zurückliegen<sup>11</sup>.

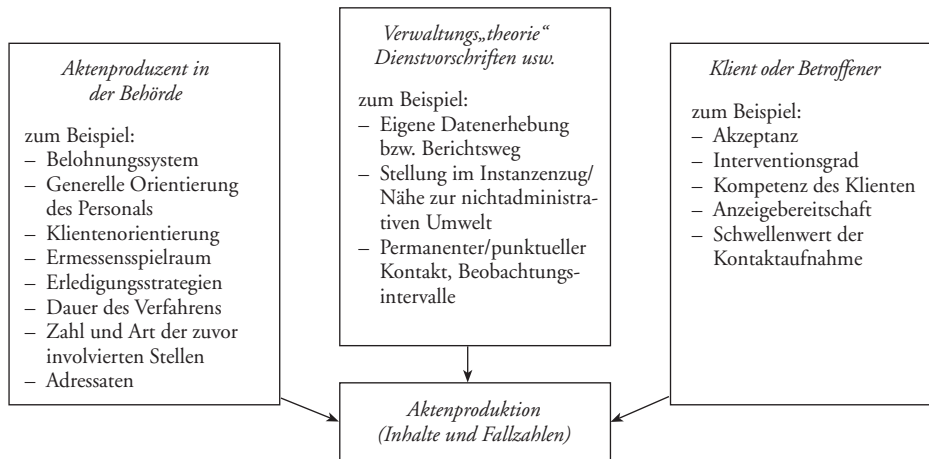
Nach Groves (1989) können in Validierungsstudien drei Herangehensweisen unterschieden werden: die rückwärtsgerichtete, die vorwärtsgerichtete und die vollständige Registerdatenüberprüfung. Die rückwärtsgerichtete Registerdatenüberprüfung geht von den im Register enthaltenen Sachverhalten aus und überprüft, inwieweit sie auch in Befragungsdaten vorkommen bzw. in welcher Hinsicht die Details von übereinstimmend angegebenen Vorkommnissen in Befragungen anders beschrieben werden. Beispielsweise befragen Cannell, Fisher und Bakker (1965) Personen, die sie den Einträgen von Krankenhäusern entnommen hatten, zu ihren Krankenhausaufenthalten. Dadurch können allerdings keine Personen in den Vergleich eingehen, die keinen Registereintrag haben, aber in einer Befragung (fälschlicherweise) ein Ereignis angeben.

Bei der vorwärtsgerichteten Registerdatenüberprüfung werden mit dem Einverständnis der Befragten relevante Register durchsucht, um deren Einträge mit den Angaben der Befragten zu vergleichen. Dies ermöglichte eine Validitätsschätzung für die Angaben einer aussagekräftigen Stichprobe. Beispielsweise gleicht Auriat (1993) die Umzüge, die ihre Befragten angegeben hatten, mit den Registern der belgischen Meldestellen ab. Ungeprüft bleiben dabei oft die Fälle, in denen Befragte keinen Sachverhalt berichten, da nach diesen auch nicht in den jeweiligen Registern gesucht wird.

---

<sup>11</sup> Aber auch hier gibt es wieder Einschränkungen. Registerdaten können für bestimmte Zeiträume schlechtere Daten enthalten, etwa für chaotische, unregelmäßige Kriegs- oder Nachkriegsperioden. Auch können sich die Verwaltungsvorschriften über die Zeit verändern.

Abbildung 3.6: Ein Datengenerierungsmodell für prozessproduzierte Daten



Quelle: Bick & Müller, 1984.

In der vollständigen Registerdatenüberprüfung (Marquis, 1978) wird eine Befragtenstichprobe gezogen und dann für alle Befragten (nicht nur für diejenigen, die entsprechende Vorkommnisse berichten) die verfügbaren Register überprüft. So können sowohl die Fälle untersucht werden, in denen die Register keinen Eintrag enthalten, wohl aber die Befragungsdaten, als auch die, in denen die Befragungsdaten ein Ereignis vermelden, nicht aber die Registerdaten. Die vollständige Registerdatenüberprüfung ist also das informativste Verfahren zur Validitätsschätzung.

### 3.3.4 Voraussetzungen für Reliabilitäts- und Validierungsstudien mit Befragungsdaten

Der Vergleich zweier Datensätze zur Reliabilitäts- bzw. Validitätsprüfung setzt voraus, dass die Datensätze in folgender Hinsicht vergleichbar sind oder vergleichbar gemacht werden können: in Bezug auf die Bezugseinheiten, die Informationen, die Vergleichszeiträume und die Datumsformate.

#### (a) Identität bzw. Vergleichbarkeit der Bezugseinheiten

Die beiden Datensätze müssen sich auf dieselben Merkmalsträger (z.B. Personen, Haushalte, Sachverhalte usw.) beziehen und diese müssen einander zugeordnet werden können. Eine Alternative wäre der Vergleich der aggregierten Angaben verschiedener Stichproben, die repräsentativ für dieselbe Grundgesamtheit gezogen wurden. Hier allerdings kann man nicht feststellen, ob Abweichungen durch den Beobachtungs- oder Nichtbeobachtungsfehler (etwa mangelnde Vergleichbarkeit der Stichproben) verursacht wurden. Außerdem ist es nicht mehr möglich, die Angaben über einen bestimmten Merkmalsträger zu vergleichen.

*(b) Identität bzw. Vergleichbarkeit der erhobenen Information*

Die Variablen in beiden Datensätzen müssen sich in ihren Merkmals- bzw. Zustandsräumen entsprechen oder zumindest ineinander überführbar sein – etwa durch die Zusammenlegung von Kategorien, wenn die Information in dem anderen Datensatz in weniger fein abgestuften Kategorien vorliegt. Beispielsweise kann der Befragte seine Bildungs- und Erwerbsaktivitäten im *British Household Panel Study* (BHPS) in eine von zwölf Kategorien einordnen, im *Sozio-ökonomischen Panel* (SOEP) stehen hingegen nur acht Kategorien zur Verfügung (<http://panel.gsoep.de/soepinfo2001>)<sup>12</sup>. Sorgfältig muss dabei auch geprüft werden, ob Konzepte bzw. Kategorien gleichen oder ähnlichen Namens identisch oder vergleichbar definiert sind. Etwa werden unter dem Begriff „Arbeitslosigkeit“ verschiedene Dinge erhoben. Manche Erhebungen verwenden die Definition der *International Labor Organization* (ILO): „(...) people that are out of work, want a job, have actively sought work in the last 4 weeks and are available to start work in the next 2 weeks OR people who are out of work, have found a job and are waiting to start it in the next 2 weeks“. Andere berücksichtigen nur Zeiträume, in denen staatliche Leistungen für Arbeitslose bezogen werden, und wieder andere überlassen es den Befragten, sich gemäß ihrer eigenen Definition als arbeitslos einzuordnen. Bei Wiederholungsbefragungen ist auch darauf zu achten, ob sich die Definitionen von Merkmals- bzw. Zustandsräumen zwischen den zu vergleichenden Erhebungen verändert haben (Jacobs, 2002; Norwood & Tanur, 1994).

*(c) Identität bzw. Vergleichbarkeit der Referenzzeiträume*

Ganz oder teilweise sich überlappende Zeiträume können relativ unproblematisch durch Trunkierung oder „Abschneiden“ aneinander angepasst werden. Dadurch entstehen aber unter Umständen an den „Rändern“ Probleme: Wenn zeitlich falsch lokalisierte Ereignisse durch das Abschneiden aus der Vergleichsperiode herausfallen, können sie in den Vergleich nicht mehr eingehen. Außerdem verkürzt sich die Dauer der abgeschnittenen Episoden.

*(d) Formale Identität bzw. Vergleichbarkeit*

Ein und dieselbe Information kann in zwei Datensätzen zwar vergleichbar enthalten, aber in unterschiedlicher Form in der Datenmatrix repräsentiert sein. Beispielsweise kann in einem Befragungsdatensatz jedes Ereignis als eine Datenzeile repräsentiert werden, während der vergleichbare Registerdatensatz Verwaltungskontakte als Datenzeilen führt. Für einen aussagekräftigen Vergleich muss es möglich sein, den einen Datensatz an das Format des anderen anzupassen bzw. beide in das gleiche Format zu überführen.

Wo Vergleichbarkeit in diesen Punkten nicht gegeben ist, kann sie unter Umständen durch geeignete Maßnahmen hergestellt werden. Wie aus den angeführten Beispielen ersichtlich wird, ist dies jedoch ein nichttrivialer Vorgang, bei dem Entscheidungen getroffen, begründet und in Hinblick auf den Zweck des Vergleichs bewertet werden müssen. Dies muss

<sup>12</sup> Im BHPS sind dies „self-employed, full-time paid employment, part-time paid employment, unemployment, retired, maternity leave, looking after family and home, full-time student/at school, long-term sick or disabled, on a government training scheme, national service/war service, something else“. Im SOEP sind dies „Schule/Studium, Lehre/Ausbildung, Wehr-, Zivildienst, Vollzeitberufstätig, Teilzeitberufstätig, Arbeitslos, Hausfrau, -mann, Ruhestand und Andere“.



auf der Grundlage von Wissen um die beiden Datenkonstruktionsprozesse und ihre Randbedingungen geschehen.

Eine große Herausforderung, speziell für die Registerdatenüberprüfung, ist die wechselseitige Zuordnung von Register- und Befragungsdaten, die sogenannte Verknüpfung oder „Record Linkage“. „If the respondent is ‚Harold Thompson‘, living at ‚123 Main Street‘, but if the record closest to a match is ‚Harold Thomas‘ at 123 Main Street, what should the researcher do? What if ‚Harold Thomas‘ lives at 123 Maynard Street?“ charakterisiert Groves (1989, S. 303) die Herausforderung. Für einzelne Ereignisse ergeben sich vergleichbare Zuordnungsschwierigkeiten: „For example, the respondent reports someone stealing the hubcaps from his car in April 1984, and the police record is for stealing of hubcaps and breaking an antenna in May 1984. Has the respondent misdated an event or reported another event?“ (Groves, 1989, S. 303) Fehler bei der Zuordnung können auch die Validitätsschätzung beeinflussen (Neter, Mayenes, & Ramanathan, 1965). Es obliegt dem Wissenschaftler, sinnvolle Zuordnungsprozeduren zu finden, die sich aus dem Wissen um die beiden Datenkonstruktionsprozesse und ihre Randbedingungen ergeben.

Die Herausforderungen bei der Analyse von Wiederholungsbefragungen sind in dieser Hinsicht geringer, da diese im Rahmen derselben Studie stattfinden und die Teilnehmer anhand unverwechselbarer Identifikationsnummern geführt werden. Auf der Ebene einzelner Vorkommnisse gibt es aber dieselben Schwierigkeiten, wenn ein Befragter in der ersten Befragung gestohlene Radkappen für April, in der zweiten hingegen gestohlene Radkappen und eine abgebrochene Antenne für Mai angibt<sup>13</sup>. Auch hier ist offensichtlich, dass diese Entscheidungen auch das Urteil über die Güte der Erinnerung zu einem späteren Zeitpunkt beeinflussen können.

Angesichts all der im vorigen Abschnitt beschriebenen Herausforderungen verwundert es nicht, dass Reliabilitäts- und Validierungsstudien meist unter suboptimalen Bedingungen durchgeführt werden. Im folgenden Abschnitt möchte ich eingehend über einige Studien berichten, die trotz all dieser Widrigkeiten die Reliabilität und Validität von Befragungsdaten betrachten und uns darüber informieren können, wie Gedächtnisprozesse die retrospektive Rekonstruktion von Lebensverläufen in standardisierten retrospektiven Befragungen beeinflussen. Ich werde dabei nur Studien berücksichtigen, die sich auf Befragungen unter authentischen Bedingungen beziehen, in denen längsschnittliche Sequenzen von Ereignissen erfasst werden und deren Befunde sich im Hinblick auf Erinnerungsfehler betrachten lassen und/oder vom Autor daraufhin betrachtet wurden. In den Abschnitten 3.4.1 und 3.4.2 werde ich allgemein über die Befunde zu Fehlern bei der retrospektiven Daten-

<sup>13</sup> Die Zuordnung einzelner Episoden oder Ereignisse aus wiederholten Befragungen muss zum Beispiel immer dann getätigt werden, wenn in Panelbefragungen die Angaben aus mehreren Wellen zu einer kontinuierlichen Sequenz zusammengeführt werden. In einer jährlichen Befragung hat ein Teilnehmer beispielsweise im Dezember 1990 angegeben, momentan arbeitslos zu sein und zwar seit Juli 1990. In der folgenden Befragung im Dezember 1991 gibt er an, momentan arbeitslos zu sein und zwar seit Juli 1991. Da der Interviewer im Moment der Erhebung zumeist keine Möglichkeit hat, die Angaben beider Wellen auf Konsistenz zu prüfen, fällt dies erst später bei der Edition und dem Zusammensetzen der Gesamtgeschichte auf. Nun kann der Wissenschaftler annehmen, dass der Befragte in der zweiten Erhebung in Wirklichkeit „Juli 1990“ meinte, und eine kontinuierliche Arbeitslosigkeitsepisode von Juli 1990 bis zur Gegenwart eintragen. Oder er nimmt an, dass für den Zeitraum von Dezember 1990 bis Juli 1991 ein anderer Zustand vorgelegen hat, über den nichts Näheres bekannt ist. Über die hier entstehenden Fehler berichten Poterba und Summers (1984) und Solga (2001).

erhebung berichten und in Abschnitt 3.4.3 gesondert über die Befunde, die Aufschluss darüber geben, ob der Fehler umso größer ausfällt, je länger die fraglichen Ereignisse zurückliegen. Um den Zweckmäßigkeitsaspekt der Datengüte zu berücksichtigen, werde ich auch die Auswirkungen der nachgewiesenen Erinnerungsfehler auf die sozialwissenschaftliche Modellbildung behandeln sowie Erhebungstechniken diskutieren, die dem Erinnerungsfehler entgegenwirken können.

### 3.4 Gedächtnisfehler in retrospektiven Ereignisdaten – Befundlage

#### 3.4.1 Fehler durch die retrospektive Rekonstruktion

Wie in Kapitel 2 beschrieben, führen die rekonstruktiven und schemageleiteten Mechanismen des autobiografischen Erinnerns dazu, dass die Erinnerungen einfacher, konventioneller und kohärenter ausfallen als die tatsächlichen Sachverhalte. Was Inzidenz und Frequenz von Ereignissen betrifft, so wird in retrospektiven Befragungen durch die Vereinfachungstendenz, das Vergessen von Episoden und das Zusammenlegen ähnlicher Episoden zu einer generischen Gesamtperiode eine zu geringe Anzahl von Episoden und Übergängen angegeben. Dies wurde in verschiedenen Validierungsstudien für retrospektive Angaben zu Partnerschaften (Klein & Fischer-Kerli, 2000), Umzügen (Auriat, 1993), Krankenhausaufenthalten (Cannell u.a., 1965), Autounfällen (Cash & Moss, 1972) und Kindern (Poulain, Riandey, & Firdion, 1991) nachgewiesen<sup>14</sup>. Verbreitung und Ausmaß der Reduktion in Episoden variieren dabei stark zwischen den erfragten Inhalten. Belli (1998) vergleicht die Anzahl der Erwerbsepisoden, die in der *Panel Study of Income Dynamics* (PSID) für eine Zweijahresperiode einmal direkt am Ende dieser Periode und erneut ein Jahr später angegeben wurden. Der Anteil der Personen mit nur einer Erwerbsepisode im Zeitraum ist in der zweiten Befragung um 8,2 Prozent größer, während der Anteil derer mit zwei oder mehr Episoden um 9,4 Prozent gesunken ist. Mit Daten des *National Longitudinal Survey of Youth 1979* (NLSY79) belegt Pierret (2001) für eine Einjahresperiode, über die die Teilnehmer ebenfalls zweimal im Abstand von einem Jahr befragt wurden, eine Verringerung um 5 Prozent für Erwerbs- und um 17 Prozent für Arbeitslosigkeitsepisoden.

Eine verringerte Episodenzahl reduziert auch die Zeit, die Personen in Zuständen eines bestimmten Typs verbringen. Wenn Ereignisdaten eine festgelegte Periode abdecken, sind die Folgen komplexer, da die durch das Vergessen von Episoden „frei werdenden“ Zeiträume „gefüllt“ werden müssen. Dies kann zum einen durch „Strecken“ einer angrenzenden Episode geschehen, zum anderen durch die Umdefinition des frei gewordenen Zeitraums in eine Episode eines anderen Typs. Wenn angrenzende Episoden gestreckt werden, führt dies zu weniger Zeit in Episoden des vergessenen Typs und zu mehr Zeit in Episoden des gestreckten Typs<sup>15</sup>.

<sup>14</sup> de Graaf und Wegener (1989) und Becker (2001) belegen zwar beim Vergleich einer Wiederholungsbefragung, dass ihre Teilnehmer in der späteren Befragung mehr Episoden berichten. In dieser Studie beträgt der abgedeckte Zeitraum allerdings bis zu 30 Jahren, die Zeit zwischen den Befragungen hingegen nur einen Monat. Insofern weist dies eher auf die generelle Instabilität autobiografischer Erinnerung hin, weniger auf Abweichungen, die durch die längere zeitliche Distanz zum Geschehen in der zweiten Erhebung verursacht werden.

<sup>15</sup> Wenn es sich um Episoden des selben Typs handelt (beispielsweise zwei Erwerbstätigkeiten), kann es dazu kommen, dass trotz abweichender Inzidenzinformation die Dauerinformation korrekt angegeben wird.

Wenn die Zeiträume umdefiniert werden, führt dies zu weniger Zeit in Episoden des ursprünglichen Typs und zu mehr Zeit in Episoden des Typs, in den umdefiniert wurde. Etwas anders sehen die Folgen beim Vergessen bzw. Einebnen von Übergängen aus: Da vorzugsweise wenig herausragende Übergänge zwischen ähnlichen Episoden eingeebnet werden, kann die Dauer in Episoden dieses Typs davon unberührt bleiben.

Für den *British Household Panel Survey* (BHPS) vergleicht Paull (2002) die Angaben zu Erwerbstätigkeit, Arbeitslosigkeit und Nichterwerbstätigkeit für einen Zeitraum von etwa einem Jahr, über den die Teilnehmer zweimal im Abstand von einem Jahr befragt wurden. Hier werden von allen Episodentypen in der späteren Befragung eine geringere Anzahl berichtet, und für alle Episodentypen haben die kurzen Episoden das höchste Risiko, abweichend wiedergegeben zu werden. Arbeitslosigkeiten werden dabei am instabilsten wiedergegeben und am massivsten reduziert. Etwa 89 Prozent der Episoden werden unverändert wiedergegeben, etwa 7 Prozent verschwinden völlig und werden durch Strecken angrenzender Episoden „eingeebnet“, und 4 Prozent der Episoden werden umdefiniert, meistens von „arbeitslos“ zu „nicht erwerbstätig“. Die Folge ist eine Verringerung der Zeiten in Arbeitslosigkeit und Erwerbstätigkeit sowie eine Zunahme der Zeiten in Nichterwerbstätigkeit. Auriant (1996) belegt beim Vergleich der Migrationsgeschichte in der „Enquête 3B.bis“ mit den Daten des belgischen Melderegisters, dass das Vergessen eines Umzugs eine erhöhte Wahrscheinlichkeit für die Fehldatierung nachfolgender Umzüge und die Überschätzung der Dauer anderer Wohnepisoden nach sich zieht.

Mathiowetz und Duncan (1988) belegen beim Vergleich von retrospektiv berichteten Arbeitslosigkeitsepisoden mit den Aufzeichnungen der jeweiligen Arbeitgeber allerdings, dass trotz massiven Vergessens von Arbeitslosigkeitsepisoden die Anzahl der dadurch nicht abgeleiteten Arbeitsstunden nicht in alarmierendem Ausmaß sinkt. Außerdem findet sich keine systematische Tendenz zum Unterschätzen dieser Arbeitsstunden, das heißt, für Dauerinformation war der Fehler nicht gerichtet und deutlich geringer als für die Inzidenz. Dieser etwas überraschende Befund lässt sich dadurch erklären, dass vor allem *sehr* kurze Arbeitslosigkeitsepisoden, die nur wenige Tage dauerten, vergessen wurden.

Weitere Sequenzeigenschaften wie Timing und Reihenfolge werden ebenfalls durch Erinnerungsprozesse beeinflusst. Das Unterschlagen von Episoden eines bestimmten Typs verlegt beispielsweise den Zeitpunkt, zu dem ein Ereignis dieses Typs zum ersten Mal auftritt, dorthin, wo die erste erinnerte Episode dieses Typs auftritt. Für eine Person, die zum Beispiel die zwei ersten Umzüge in den späten 1970er Jahren vergisst, wird der dritte Umzug in den 1990er Jahren als das erste Migrationsereignis überhaupt geführt. Dadurch könnten irrtümlicherweise für frühere historische Perioden oder frühere Lebensalter angenommen werden, dass Ereignisse dieses Typs dort (noch) nicht oder nicht so häufig auftraten. Hierzu sind mir keine systematischen Befunde bekannt (einige Hinweise finden sich in den in Abschnitt 3.4.2 berichteten Studien zu den Folgen des Erinnerungsfehlers für inhaltliche Ergebnisse).

Was die Abhängigkeit des Fehlers vom tatsächlichen Sachverhalt anbelangt, so betrifft das Vergessen vor allem kurze und untypische, nicht der (subjektiven oder allgemeinen) Norm entsprechende Ereignisse, die bei den Befragten nicht stabil und im Einklang mit den Konzepten des Wissenschaftlers repräsentiert sind. Dies hat drei Folgen: Erstens werden innerhalb bestimmter Ereigniskategorien (etwa Erwerbstätigkeiten) die kurzen und untypischen schlechter erinnert. Dies belegen unter anderem Mathiowetz und Duncan (1988) für Arbeitslosigkeitsepisoden, Klein und Fischer-Kerli (2000) für Partnerschaften und Courgeau

(1991) für Wohnepisoden. Zweitens werden bestimmte Ereigniskategorien über alle Personen hinweg schlechter erinnert als andere. Beispielsweise werden Arbeitslosigkeitsepisoden deutlich häufiger vergessen als Erwerbstätigkeitsepisoden (Dex & McCulloch, 1997; Elias, 1997; Paull, 2002, siehe Näheres zum Thema Arbeitslosigkeiten in Abschnitt 5.2.2). Zum Dritten sind dadurch ganze Lebensläufe schwerer betroffen als andere: Komplexe und unkonventionelle Ereignissequenzen mit zahlreichen Wechseln, kurzen Episoden und unscharfen Übergängen werden zu späteren Zeitpunkten deutlich inkonsistenter rekonstruiert als stabile und wenig abwechslungsreiche Verläufe, die sich unproblematisch in die vorgegebenen Zustandsräume fügen. Beispielsweise steigert eine hohe Anzahl zu erinnernder Episoden und Übergänge die Wahrscheinlichkeit, inkonsistente Angaben zu machen: Pierret (2001) zeigt in der bereits beschriebenen Analyse einer Wiederholungsbefragung, dass diejenigen, die in der ersten Befragung nur eine Episode berichtet hatten, dieselbe in über 97 Prozent der Fälle wieder angeben, während es von den Personen mit drei oder mehr Episoden nur noch 73 Prozent sind, die die Episoden in der zweiten Befragung genauso berichten. Ähnliches belegen Klein und Fischer-Kerli (2000) bei einem Vergleich der sechs Jahre auseinander liegenden Angaben zur Partnerschaftsgeschichte im „Familiensurvey“ des Deutschen Jugendinstituts und Tölke (1980) bei einem Vergleich der zehn Jahre auseinander liegenden Berufsgeschichten in der „Konstanzer Mobilitätsstudie“. Auriat (1996) hingegen konnte keinen Einfluss der Anzahl der zu berichtenden Umzüge auf das Risiko, einen oder mehrere zu vergessen, belegen.

Einige Studien belegen auch, dass persönliche Merkmale der Befragten Einfluss auf die Güte retrospektiver Daten haben. Beispielsweise erinnern sich Frauen besser als Männer an Umzüge, das Datum von Eheschließungen und Kindsgeburten (Auriat, 1993; Klijzing & Prophet, 1998), aber schlechter an Ausbildungen und Erwerbstätigkeiten (de Graaf & Wegener, 1989; Dex & McCulloch, 1997; Elias, 1997; Paull, 2002; Thélot, 1990). Letzteres wird damit begründet, dass Ausbildungs- und Berufsverläufe von Frauen stärker von kürzeren und zahlreicheren Episoden, geringfügigen Beschäftigungen und weniger eindeutigen Wechseln zwischen verschiedenen Arten von Nichterwerbstätigkeit gekennzeichnet sind. Becker (2001) und de Graaf und Wegener (1989) zeigten, dass Beamte und Angestellte im öffentlichen Dienst deutlich unzuverlässigere Angaben zu ihren Berufsverläufen machten, da Karrieren im Staatsdienst oft aus weniger eindeutig abgrenzbaren Episoden bestehen. Vielfach lassen sich Unterschiede in der Erinnerungsgüte zwischen Personengruppen also dadurch erklären, dass eine der Gruppen häufiger fehleranfällige Lebensverläufe aufweist. Das schlechtere Abschneiden der Männer bei Angaben aus den Bereichen Familie und Migration hingegen lässt sich nicht durch Unterschiede in der zu erinnernden Information erklären, besonders nicht in der Studie von Auriat (1993), in der die verglichenen Männer und Frauen jeweils Ehepartner waren.

Manche Studien belegen, dass ältere Befragte sich schlechter erinnern, besonders ab etwa dem 60. Lebensjahr (Elias, 1997; Klijzing & Prophet, 1998; Paull, 2002); andere hingegen finden keinen Zusammenhang zwischen Alter und Erinnerungsgüte (Klein & Fischer-Kerli, 2000). Effekte des Alters sind schwer zu interpretieren, da Alter mit einer Reihe ebenfalls erinnerungsrelevanter Faktoren konfundiert ist. Interaktionseffekte von Alter und Geschlecht treten ebenfalls häufig auf: Beispielsweise erinnern sich junge Frauen am allerschlechtesten an ihre Erwerbs- und besonders Arbeitslosigkeitsepisoden (Elias, 1997; Paull, 2002).

Zusammenfassend lässt sich also sagen, dass in den besprochenen Studien die Folgen von autobiografischen Gedächtnisprozessen vor allem für die Sequenzeigenschaften Inzidenz bzw.

Frequenz und Dauer von Episoden betrachtet wurden. Die gefundenen Abweichungen deuten darauf hin, dass Fehler verbreitet sind und auf Vereinfachung und Konventionalisierung durch Reduktion in Episoden und Übergängen hinauslaufen. Abweichungen der Dauerinformationen sind ebenfalls verbreitet und können beträchtlich sein. Außerdem treten sie häufig in einer bestimmte Richtung auf. Alle Fehlereigenschaften sind in ausgesprochen starkem Maße abhängig von den zu erfassenden tatsächlichen Sachverhalten: Für Ereignisse mit bestimmten Eigenschaften (kurz, unwichtig), bestimmte Ereignistypen (kurze) und bestimmte Ereignissequenzen (komplexe, unkonventionelle, abwechslungsreiche) treten Fehler deutlich häufiger, größer und gerichteter auf als für andere und führen zu einer stärkeren Vereinfachung, Konventionalisierung und Glättung der erinnerten Lebensverläufe.

### 3.4.2 Folgen für die inhaltlichen Auswertungen

In einigen Studien wurde der Effekt der gedächtnisbedingten Abweichungen auf die Ergebnisse quantitativer sozioökonomischer Analysen betrachtet, indem dieselben regressionsanalytischen Auswertungen einmal mit Daten aus früheren und ein zweites Mal mit Daten aus der späteren Befragung durchgeführt wurden. Courgeau (1991) vergleicht die Angaben zur Migrationsgeschichte in der „Enquête 3B.bis“ mit den Daten des belgischen Melderegisters und findet keine nennenswerten Unterschiede zwischen den Datenquellen für die Dauer des Aufenthalts in der ersten Wohnung nach der Eheschließung und deren Abhängigkeit von verschiedenen Personenfaktoren und Ereignissen in anderen Lebensbereichen (z.B. die Geburt eines Kindes).

In einem Vergleich von Wiederholungsbefragungen berechnen Klein und Fischer-Kerli (2000) die Survivorfunktionen für den Übergang in die Ehe und für die Ehedauer der verheirateten Teilnehmer auf der Basis von sechs Jahre auseinander liegenden Angaben. Die ermittelten Funktionen unterschieden sich nicht nachweisbar zwischen den Datenquellen.

Ebenfalls bei einem Vergleich von Wiederholungsbefragungen prüft Paull (2002) die Zusammenhänge zwischen der Höhe des aktuellen Einkommens einerseits und Berufserfahrung und Beschäftigungsdauer andererseits. Sie errechnet dabei die Berufserfahrung aus der Anzahl der Jahre, die eine Person in ihrem Leben erwerbstätig war, und die Beschäftigungsdauer aus der Anzahl der Jahre beim selben Arbeitgeber. Die Zusammenhänge fallen bei Verwendung der Retrospektivdaten schwächer aus, wobei diese Reduktion für Frauen deutlicher ausgeprägt ist<sup>16</sup>. Dies kann darauf zurückgeführt werden, dass durch fehlerhafte Rekonstruktionsprozesse in den retrospektiv erhobenen Daten einerseits das Vergessen von Episoden die Dauer in Erwerbstätigkeit unterschätzt wird und andererseits durch das „Ein-ebnen“ von Übergängen die Kontinuität von Beschäftigung überschätzt wird, wodurch die Varianz eingeschränkt und der Zusammenhang gelockert wird.

Ebenfalls bei einem Vergleich von Wiederholungsbefragungen betrachtet Pierret (2001) das Risiko für Erwerbstätige, innerhalb eines Einjahreszeitraums Leistungen der staatlichen

<sup>16</sup> Allerdings enthalten Ereignisgeschichten, die aus den Angaben mehrerer Panelwellen zusammengesetzt werden, tendenziell zu viele Episoden und Übergänge, sodass Diskrepanzen nicht ausschließlich auf *under-reporting* durch Erinnerungsmängel in den retrospektiven Daten zurückzuführen sein müssen (Poterba & Summers, 1984; Solga, 2001)

Wohlfahrt zu beziehen (Inzidenz), und den Zeitpunkt, zu dem dieser Wohlfahrtsbezug zum ersten Mal eintrat (Timing)<sup>17</sup>. Dies setzt er in Bezug zu verschiedenen Eigenschaften der Personen (z.B. Geschlecht, Ethnie, Familienstand, Wohngegend) und früheren Ereignissen (z.B. Dauer früherer Erwerbstätigkeiten und Arbeitslosigkeitsphasen). Manche der Faktoren üben nur bei Verwendung der Paneldaten und nicht bei Verwendung der retrospektiven Daten einen signifikanten Einfluss darauf aus, ob bzw. wann Leistungen der staatlichen Wohlfahrt in Anspruch genommen wurden. Dabei reagiert das Modell, das nicht nur die Wahrscheinlichkeit, sondern auch den Zeitpunkt des Eintretens eines Übergangs betrachtete, empfindlicher auf die Unterschiede in den Datenquellen. Der Autor führt dies darauf zurück, dass das Vergessen hauptsächlich kurze Episoden von Leistungsbezug und Personen, die üblicherweise keine staatlichen Wohlfahrtsleistungen beziehen, betrifft. Dadurch fehlen diese *light users* in der Gruppe derer, die im abgedeckten Zeitraum Wohlfahrtsleistungen beziehen, und die mit längerem, häufigerem und „gewohnheitsmäßigem“ Leistungsbezug einhergehenden Faktoren gewinnen in den Modellen an Gewicht.

Insgesamt können also die durch Gedächtnisprozesse verursachten Datenfehler auch die Parameter von sozioökonomischen Modellschätzungen in ihrer Signifikanz, Höhe und Wirkrichtung beeinflussen. Am labilsten sind dabei Variablen, in denen Information über zeitliche An- und Einordnung von Ereignissen enthalten ist. Schömann (1994) allerdings, der das Einstiegsgehalt der ersten Erwerbstätigkeit mit Faktoren wie Bildungsniveau, Kohortengröße und Firmengröße in Bezug setzt, berichtet, dass sich die Modellkoeffizienten nur sehr geringfügig ändern und dass die Anreicherung des Modells mit zusätzlichen Erklärungsvariablen wesentlich größere Veränderungen nach sich zieht<sup>18</sup>.

Die bisher besprochenen Befunde beziehen sich auf Fehler, die prinzipiell durch die Retrospektivität der Angaben verursacht werden. Damit ist noch nicht gesagt, dass oder in welcher Weise Fehler und Abweichungen gravierender ausfallen, je länger die Ereignisse zurückliegen. Im folgenden Abschnitt möchte ich über Befunde berichten, die im Hinblick auf das Ausmaß der retrospektiven Distanz interpretiert werden können und einige grundsätzliche Überlegungen zu diesem Thema anstellen.

### 3.4.3 Fehler und die Länge der retrospektiven Distanz – Befunde und grundsätzliche Überlegungen

Die bisher berichteten Befunde informieren uns zwar über Ausmaß und Art von Abweichungen, die durch die Retrospektivität bzw. die Zunahme der retrospektiven Distanz verursacht werden, verraten aber noch nichts über den genauen Zusammenhang der Länge der retrospektiven Distanz mit der Zunahme des Fehlers. Dies lässt sich an Ereignisdaten allerdings

<sup>17</sup> Dabei handelt es sich zum einen um finanzielle Unterstützungsleistungen, die die Bundesstaaten der Vereinigten Staaten von Amerika im Rahmen des *Aid to Families with Dependent Children*-Programms an bedürftige Familien mit minderjährigen Kindern auszahlen, und zum anderen um Kaufgutscheine für Lebensmittel, die ebenfalls auf bundesstaatlicher Ebene im Rahmen des *Food Stamps*-Programms an Bedürftige ausgegeben werden.

<sup>18</sup> Allerdings liegen dieser Analyse die in Fußnote 14 beschriebenen Daten zu Grunde, die wegen des Zwischenerhebungszeitraums von nur einem Monat nicht unbedingt aufschlussreich für die Auswirkungen von spezifisch erinnerungsbedingter Abweichungen der Angaben sind.

gut überprüfen: Retrospektiv erhobene Ereignissequenzen decken immer eine ausgedehnte Periode ab. Wenn Erinnerungsfehler mit fortschreitender Zeit zu einer kontinuierlichen Verschlechterung der Angaben führen, sollten die wiederholten Angaben für spätere Ereignisse in dieser Periode auf allen Fehlerdimensionen besser übereinstimmen als für länger zurückliegende.

Dass Fehler für länger zurückliegende Zeiträume verbreiteter sind, belegen zum Beispiel Klein und Fischer-Kerli (2000) für Angaben zur Partnerschaftsgeschichte und Cannell u.a. (1965) und Cash und Moss (1972) für Krankenhausaufenthalte. Hingegen findet sich bei Auriat (1993) kein Einfluss der Länge der verstrichenen Zeit auf die Verbreitung und das Ausmaß der Fehler beim zeitlichen Einordnen von Umzügen, Hochzeiten und Geburtstagen von Kindern. Auch für Arbeitslosigkeitsepisoden im „BHPS“ vergleicht Elias (1997) die Angaben von 1990 und 1991 über das jeweils vergangene Jahr mit den Angaben einer Wiederholungsbefragung 1992, die beide Jahre abdeckt, und stellt für das länger zurückliegende Jahr 1990 und das kürzer zurückliegende Jahr 1991 gleiche Übereinstimmungsraten fest.

Es ist allerdings wichtig, genau zu klären, auf welchem Wege die verlängerte retrospektive Distanz einen Einfluss auf die Übereinstimmung bzw. die Fehlerhaftigkeit ausübt. In realistischen Befragungssituationen gehen eine ganze Reihe von Faktoren, die die Erinnerungsgüte beeinflussen können, mit der verstreichenden Zeit einher. Deren jeweiliger Einfluss lässt sich in Reliabilitäts- und Validierungsstudien – anders als in kontrollierten gedächtnispsychologischen Studien – nicht oder nur selten voneinander isoliert betrachten.

Erstens ergibt sich bei Ereignisdaten aus einer verlängerten retrospektiven Distanz ein verlängerter Zeitraum, über den Bericht zu erstatten ist und damit eine erhöhte Menge von Information, die rekonstruiert werden muss, was sowohl die kognitive Anforderung steigert als auch das Risiko, dass Ereignisse miteinander verwechselt werden. Pierret (2001) zeigt in seiner Wiederholungsbefragung, dass die Übereinstimmung für ein Jahr, über das mit identischer retrospektiver Distanz berichtet wurde, niedriger ausfiel, wenn dieses Jahr nicht das einzige abgefragte, sondern Teil einer längeren (zweijährigen) retrospektiven Periode war.

Zweitens ist die retrospektive Distanz mit dem Alter der befragten Person beim Erleben eines Ereignisses konfundiert: Ein und dieselbe retrospektive Distanz bedeutet für manche Befragte, dass sie sich an das junge Erwachsenenalter erinnern, für manche, dass sie sich an ihre mittleren Jahre erinnern. Zum einen kann dadurch ein Reminiszenzphänomen (siehe Abschnitt 2.6.2) auftreten, wodurch Ereignisse aus dem Jugend- und frühen Erwachsenenalter besser erinnert werden, gleichgültig, wie lange dieser Lebensabschnitt für die Erinnernden zurückliegt. Zum anderen gehen bestimmte Lebensalter mit schwerer rekonstruierbaren Ereignissen und Verläufen einher. Beispielsweise bringt die Phase des Berufseintritts im jungen Erwachsenenalter häufig eine Reihe von raschen Wechseln mit sich, die im Vergleich zu Wechseln der mittleren Lebensjahre weniger zielgerichtet und weniger folgenschwer und somit ungeachtet der seither verstrichenen Zeit schwieriger zu rekonstruieren sind. Ähnliches mag für frühe Partnerschaften jüngerer Menschen gelten. Andererseits sollten gerade die *first time events* (die erste „richtige“ Erwerbstätigkeit, die erste Wohnung) in der Struktur des autobiografischen Gedächtnisses besonders gut verankert sein, da sie einen folgenreichen Übergang markieren und für das Selbst von zentraler Bedeutung sind (Robinson, 1992). So belegten de Graaf und Wegener (1989), dass der Beginn des ersten Erwerbsverhältnisses in einer Wiederholungsbefragung nur geringfügig abweichender datiert wurde als der Beginn der gegenwärtigen Erwerbstätigkeit, während dazwischenliegende Erwerbsverhältnisse deutlich schlechter datiert wurden.

Drittens können unterschiedlich lange retrospektive Intervalle auch verschiedene Perioden betreffen, die (für alle Befragten) erschwerte Erinnerungsbedingungen schaffen. Dies können historische Perioden sein, wenn durch Kriegs- und Nachkriegsbedingungen, Wirtschafts- und politische Krisen oder andere Umwälzungen individuelle Lebensverläufe chaotisch, ziellos und außerhalb geregelter Institutionen verlaufen (Brückner, 1993). Auch aus anderen Gründen können Zeiträume für alle Erinnernden schwierigere Bedingungen mit sich bringen. Etwa stellen Mathiowetz und Duncan (1988) in einem Vergleich von retrospektiv berichteten Arbeitslosigkeitsepisoden mit den Aufzeichnungen der jeweiligen Arbeitgeber über einen Referenzzeitraum von zwei Jahren fest, dass Arbeitslosigkeitsphasen aus unbekanntem Gründen im weiter zurückliegenden Jahr wesentlich häufiger aufgetreten und kürzer und unregelmäßiger ausgefallen waren. Nach der statistischen Kontrolle für Episodenlänge war die Übereinstimmung für das weiter zurückliegende Jahr nicht mehr schlechter als für das Jahr direkt vor der Befragung.

Viertens besteht die Möglichkeit, dass sich durch individuelle oder historische Entwicklungen und Erschütterungen auch die Wahrnehmung der eigenen Biografie so verändert, dass die zur Selbstbeschreibung angebotenen Zustände zu einem späteren Zeitpunkt anders angewendet werden. Dabei werden die abweichenden Interpretationen vermutlich von den Anforderungen der aktuellen Lebenslogik mit gestaltet: Beispielsweise kann sich eine Partnerschaft, die zu einem früheren Zeitpunkt als ausgenommen wichtig erlebt und berichtet wurde, in der Wahrnehmung einer später anderweitig verheirateten Person als unwichtige und deshalb nicht zur Kategorie ernsthafter Partnerschaften gehörende Episode darstellen.

Alles in allem bestimmen also mit zunehmender retrospektiver Distanz zahlreiche Faktoren gemeinsam, ob Angaben übereinstimmend ausfallen. Viele dieser mit der retrospektiven Distanz assoziierten Faktoren deuten dabei in dieselbe Richtung und führen dazu, dass mit längeren retrospektiven Distanzen die Rekonstruktion erschwert wird und Fehler zunehmen. Deshalb muss man tatsächlich annehmen, dass in retrospektiven Befragungen eine zusätzliche Fehlerkomponente enthalten ist, die eine Abweichung der Daten von der Realität in einer ganz charakteristischen Form herstellt. Dieser Einfluss fällt allerdings für verschiedene Inhalte sehr unterschiedlich aus. Auch weiß man nicht sicher, bei welcher retrospektiven Distanz man mit welcher Zunahme an Fehlerhaftigkeit der einen oder anderen Art rechnen muss, ob diese Zunahme sich stetig fortsetzt und ab welchem Punkt für welche Art von Information und welchen Fehleraspekt keine weitere Zunahme mehr erfolgt.

#### 3.4.4 Validitätssteigernde Erhebungsbedingungen und -techniken

Unabhängig von allen erinnerungsrelevanten Eigenschaften der Information, der Person oder der Länge der retrospektiven Distanz gelingt das Erinnern den Befragten besser, wenn die Erinnerungssituation einerseits den Stärken des autobiografischen Erinnerns Raum zur Entfaltung gibt und andererseits seinen Schwächen unterstützend entgegenwirkt. Ersteres bedeutet, dass die Gestaltung der Erhebungssituation deshalb die flexible Nutzung der individuellen horizontalen, hierarchischen und chronologischen Rekonstruktionspfade und den Einsatz der ebenso individuellen komplementären Strategien zur inferenziellen Rekonstruktion von Zeit- und Datumsinformationen ermöglichen sollte. Dadurch können die Befragten eigenständig unterstützende Erinnerungs-Hinweisreize generieren und auch schwer zu erinnernde Infor-



mation rekonstruieren und absichern. Andererseits sollte Hilfestellung bei der Rekonstruktion schwer zu erinnernder Inhalte wie kurzen, unwichtigen und seriellen Ereignissen geboten und Möglichkeiten zur Absicherung und Überprüfung von Datumsangaben an Grenzsteinereignissen gegeben werden.

Obgleich standardisierte Befragungen den idiosynkratischen Strukturen des autobiografischen Erinnerns nur begrenzt entgegenkommen können, haben sich eine Reihe von Maßnahmen als wirksame Gedächtnishilfen erwiesen.

*(a) Chronologische Rekonstruktion*

Für Ereignisabfolgen, die wie Schul- oder Erwerbskarrieren eine innere zeitliche Logik aufweisen, ist es am vorteilhaftesten, dieser Richtung zu folgen und vom ersten Ereignis ausgehend bis zur Gegenwart voranzuschreiten. Ist dies nicht der Fall – zum Beispiel für Viktimisierungen oder Krankenhausaufenthalte –, hat sich zumindest für kürzere Zeiträume die umgekehrt chronologische Ordnung besser bewährt (Loftus & Fathi, 1985; Means & Loftus, 1991).

*(b) Einsatz von Referenzereignissen*

Die Begrenzung des abgedeckten Zeitraums durch ein unverwechselbares und zeitlich eindeutig lokalisiertes Referenzereignis erleichtert den Befragten zu entscheiden, ob ein Ereignis in die abgefragte Periode fällt. Loftus und Marburger (1983) verwendeten den verheerenden Ausbruch eines in der Nähe gelegenen Vulkans. In der *Deutschen Lebensverlaufsstudie* (LV) wurden die Verläufe ab dem Mauerfall erhoben (Solga, 1996). Auch für die Rekonstruktion von Kalenderdaten können Referenzereignisse innerhalb der abgefragten Periode Ankerpunkte bieten (Belli, 1998; Bird, Born, & Erzberger, 2000; Loftus & Marburger, 1983).

*(c) Individuelle Kontextualisierung mit anderen Lebensbereichen*

Während eine bereichsspezifische Abfrage der thematischen Struktur des autobiografischen Gedächtnisses entgegenkommt, sind thematische Bereiche wie Arbeiten, Wohnen und Partnerschaft im Gedächtnis auch miteinander verbunden (Barsalou, 1988; Conway, 1996). Wenn andere Bereiche als Kontexte aktiviert werden, werden dem Befragten deshalb zahlreiche individuelle Abrufhinweise und Referenzereignisse zugänglich gemacht, anhand derer er Rekonstruktionen und ihre zeitliche Platzierung vornehmen und absichern kann.

Eine Technik, in der all diese unterstützenden Maßnahmen verwirklicht werden, ist der biografische Kalender (auch *Life History* oder *Event History Calendar*). Abbildung 3.7 zeigt einen Ausschnitt aus einem Kalender, der in der *Panel Study of Income Dynamics* zur Erhebung der Erwerbsbiographie verwendet wurde (Belli u.a., 2001). In einer großen Matrix werden auf einer Achse alle interessierenden Lebensbereiche abgetragen – zum Beispiel Wohn-, Erwerbs- und Krankheitsgeschichte. Rechtwinklig dazu wird die chronologische Zeit in Tagen, Wochen, Monaten oder Jahren abgetragen. Innerhalb der Zellen des Rasters können nun vom Interviewer Anfang und Ende von Episoden grafisch festgehalten werden (Balan, Browning, Jelin, & Litzler, 1969; Belli, 1998; Belli u.a., 1999; Caspi u.a., 1996; Freedman, Thornton, Camburn, Alwin, & Young-DeMarco, 1988; Marcoux, 1997). Solche Instrumente visualisieren außerdem die interdependente Episodenstruktur des Lebens, sodass Inkonsistenzen und Lücken dem Befragten selbst oder dem Interviewer eher auffallen. Darüber ist es möglich, in

Abbildung 3.7: Biografischer Kalender zur Erhebung der Lebensverlaufsdaten

		Winter 1996			Spring 1996			Summer 1996			Fall 1996		
		January	February	March	April	May	June	July	August	September	October	November	December
1st					27th Memorial			4th Independence			2nd Labor Day		
New Years					L A N D M A R K			E V E N T S			28th Thanksgiving		
		[2 1/4 years ago]			[2 years ago]			[1 3/4 years ago]			[1 1/2 years ago]		
					S T R E E T			C I T Y			S T A T E		
					S T R E E T			C I T Y			S T A T E		
I		Winter 1996			Spring 1996			Summer 1996			Fall 1996		
#		January	February	March	April	May	June	July	August	September	October	November	December
1		X											
2		2	2	2	2	2	2	2	2	2	2	2	2
3		3	3	3	3	3	3	3	3	3	3	3	3
4		4	4	4	4	4	4	4	4	4	4	4	4
5		5	5	5	5	5	5	5	5	5	5	5	5
6		6	6	6	6	6	6	6	6	6	6	6	6
7		7	7	7	7	7	7	7	7	7	7	7	7
8		8	8	8	8	8	8	8	8	8	8	8	8
9		9	9	9	9	9	9	9	9	9	9	9	9
10		10	10	10	10	10	10	10	10	10	10	10	10
11		11	11	11	11	11	11	11	11	11	11	11	11
12		12	12	12	12	12	12	12	12	12	12	12	12
J		Winter 1996			Spring 1996			Summer 1996			Fall 1996		
#		January	February	March	April	May	June	July	August	September	October	November	December
		[2 1/4 years ago]			[2 years ago]			[1 3/4 years ago]			[1 1/2 years ago]		
1		1	1	1	1	1	1	1	1	1	1	1	1
2		2	2	2	2	2	2	2	2	2	2	2	2
								E M P L O Y			M E N T		

Ausschnitt aus einem in der *Panel Study of Income Dynamics* verwendeten biografischen Kalender zur Erhebung der Erwerbsbiografie.

Quelle: Originalfragebogen.

einem für standardisierte Interviews ungewöhnlichem Ausmaß auf die individuellen Rekonstruktionsstrategien und -präferenzen der Befragten einzugehen, da man gut geschulten Interviewern freie Hand über die Reihenfolge der Rekonstruktion lassen kann, solange das Raster nur vollständig ausgefüllt wird.

Belli u.a. (2001) führten die in Abschnitt 3.4.1 beschriebene Wiederholungsbefragung für einen Teil der Befragten mit einem biografischen Kalender durch. Während die mit der herkömmlichen Fragenliste befragten Teilnehmer in der Wiederholungsbefragung deutlich weniger Episoden berichtet hatten, stimmte in der Gruppe der mit dem Kalender befragten Teilnehmer der Anteil der Befragten mit einer bzw. zwei Episoden während des abgedeckten Zeitraums nahezu perfekt überein. Noch ungeklärt ist, ob einzelne Aspekte des Kalenders

für diese Verbesserung verantwortlich sind oder ob die Gesamtheit der unterstützenden Faktoren zusammenwirken muss. Da der Kalender in telefonischen Befragungen zu einer Verbesserung der Übereinstimmung führte (Belli u.a., 2001), lässt sich vermuten, dass die Visualisierung eher den Interviewer dabei unterstützt, die Angaben auf Stimmigkeit zu überprüfen.

#### 3.4.5 Zusammenfassung: Gedächtnisfehler in Retrospektivbefragungen

Die in der psychologischen Gedächtnisforschung nachgewiesene Neigung, den eigenen Lebenslauf in der Rückschau durch selektives Vergessen und Umdeuten von Episoden und Übergängen als kontinuierlicher, kohärenter und konventioneller zu rekonstruieren, als es den tatsächlichen Gegebenheiten entspricht, macht sich auch bei der retrospektiven Erhebung sozialwissenschaftlicher Lebensverlaufsdaten in standardisierten Interviews bemerkbar. Das Vergessen von (kurzen) Episoden und das Einebnen von unscharfen Übergängen zwischen seriellen Ereignissen führen dazu, dass Lebensläufe in der Rückschau weniger wechselhaft erscheinen, als sie tatsächlich waren. Der Nachteil für untypische Erinnerungen, die der aktuellen subjektiven und normativen Entwicklungslogik des Befragten nicht entsprechen, führt zu konventionelleren und folgerichtigeren Sequenzen.

Besonders davon betroffen sind ereignisreiche, wechselhafte, wenig folgerichtige und unkonventionelle Lebensläufe; der Fehler tritt also abhängig vom tatsächlichen Sachverhalt auf und führt zu einer Einschränkung der Varianz. Sozialwissenschaftlich relevante Gruppen wie Geschlechts- oder Berufsgruppen sind davon stärker oder schwächer betroffen, was vielfach darauf zurückzuführen ist, dass diese sich hinsichtlich der zu erinnernden Lebensläufe unterscheiden. Teilweise schlagen sich diese Unterschiede auf die Ergebnisse sozialwissenschaftlicher Erklärungsmodelle nieder, besonders, wenn in den abhängigen oder unabhängigen Variablen zeitliche oder Verlaufsaspekte der Lebensläufe kodiert sind. Erhebungsprozeduren, die die Mechanismen des Erinnerns berücksichtigen, können allerdings den Nachteilen entgegenwirken und dadurch Erinnerungsfehler reduzieren oder sogar ganz ausschalten.

Wie genau der Zusammenhang zwischen verstreichender Zeit und Erinnerungsgüte aussieht, ist nicht eindeutig geklärt. Zum einen hat dies praktische Gründe: Viele Reliabilitäts- und Validitätsstudien beziehen sich auf relativ kurze Zeiträume von zwei bis drei Jahren, etwa auf die ein- bis zweijährigen Zwischenerhebungszeiträume von Panelbefragungen. Aber auch die Tatsache, dass gerade für lange Zeiträume eine Vielzahl von erinnerungsrelevanten Faktoren mit der retrospektiven Distanz einhergehen, die in dieselbe oder eine andere Richtung zeigen und für verschiedene Befragte unterschiedlich wirksam werden, kompliziert derlei Vergleiche und erschwert ihre Interpretation.

Auch über die Rekonstruktion von Ereignissequenzen und den Fehlern, die dabei entstehen, ist noch zu wenig bekannt. Bei diesen komplexen Daten ist es eine besondere Herausforderung, Fehler aussagekräftig zu beschreiben und auf einzelne kognitive Prozesse zurückzuführen. Oft werden Sequenzeigenschaften verglichen, also übergeordnete Information über Inzidenz, Frequenz und Dauern. Gerade die Befunde von Paull (2002) (siehe Abschnitt 3.4.1) belegen, dass die Abweichungen zwischen Sequenzen nicht umstandslos als Vergessen und Fehlerinnern beschrieben werden können, es stellen sich vielmehr Fragen wie diese: Was geschieht mit Zeiträumen, die in Erwerbstätigkeit verbracht wurden, was mit denen, die in Ar-

beitslosigkeit verbracht wurden? Wird eine der angrenzenden Episoden gestreckt und wenn ja, welche? Oder wird der Zeitraum als eine Episode anderen Typs berichtet?

Diese Aspekte möchte ich deshalb in den empirischen Analysen in den Kapiteln 5 und 6 untersuchen. Im folgenden Absatz werde ich die *Deutsche Lebensverlaufstudie* (LV), die mir dies ermöglicht, eingehend und unter erinnerungsrelevanten Gesichtspunkten beschreiben.

### 3.5 Die *Deutsche Lebensverlaufstudie* (LV)

#### 3.5.1 Ziele und Konzepte der *Deutschen Lebensverlaufstudie* (LV)

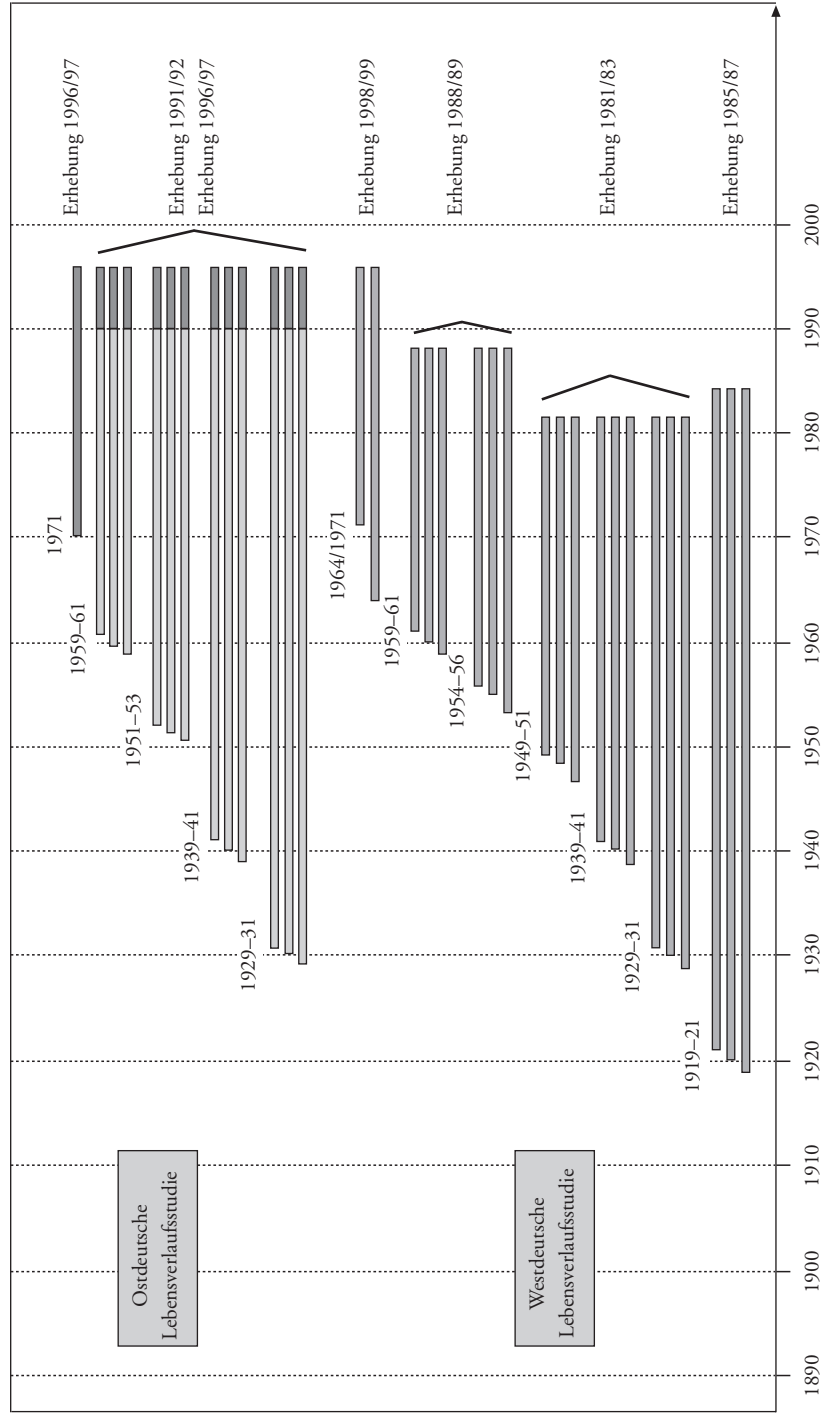
Das Projekt „Lebensverläufe und gesellschaftlicher Wandel“ wird unter der Leitung von Karl Ulrich Mayer seit 1979 und seit 1983 am Forschungsbereich „Bildung, Arbeit und gesellschaftliche Entwicklung“ des Max-Planck-Instituts für Bildungsforschung durchgeführt. In der LV wurden lebensgeschichtliche Individualdaten als Grundlage für die beschreibende und erklärende Analyse von inter- und intragenerationaler sozialer Mobilität, der Interaktion von beruflicher und Familienentwicklung und der diesbezüglichen zeit- und kohortenspezifischen Unterschiede erhoben. Seit 1981 wurden mehr als 10.000 national repräsentativ ausgewählte Teilnehmer aus acht Geburtskohorten aus Ost- und Westdeutschland retrospektiv befragt (siehe Abb. 3.8), für einen Überblick über die Erhebung und die Daten siehe Solga (1996) und Wagner (1996b).

Individuelle Lebensverläufe werden in diesem Forschungsprogramm als quantifizierbare und intersubjektiv vergleichbare Abfolgen von Ereignissen oder Zuständen begriffen, die sich entlang der Zeitachse in mehreren unterschiedlichen, aber aufeinander bezogenen Lebensbereichen entfalten. Die Zustandsräume reflektieren die durch eine Gesellschaft mehr oder weniger institutionell geregelten Pfade, wie zum Beispiel Berufskarriere oder Wohnungsgeschichte; Übergänge oder Transitionen markieren das Ende der einen und den Beginn einer anderen Episode. Dabei interessieren weniger die einzelnen, isolierten Ereignisse, sondern vielmehr der Verlaufscharakter des Lebenslaufs. Die Eigenschaften der Sequenzen wie Inzidenz, Dauer, Timing und zeitliche Verortung von Ereignissen im Lebensverlauf werden als abhängige und unabhängige Variablen in den Analysen eingesetzt und mit Merkmalen der Person, ihrer engen und weiteren sozialen Kontexte sowie anderen Eigenschaften der Ereignissequenz selbst in Beziehung gesetzt. Für den Einsatz von Analysetechniken wie der Ereignisanalyse (*Event History Analysis*), die der komplexen Rolle verschiedener Zeitdimensionen für Lebensverläufe gerecht werden, müssen die Episoden lückenlos erfasst und die Übergänge auf einer angemessenen Zeitskala datiert werden (Blossfeld & Rohwer, 1995; Blossfeld, Hamerle, & Mayer, 1991; Mayer & Huinink, 1990; Rohwer, 1994).

#### 3.5.2 Erinnerungsrelevante Erhebungsumstände in der LV

Die Befragten wurden aufgefordert, ihre Ereignissequenzen innerhalb getrennter Lebensbereiche ohne Unterbrechung oder Überlappung in chronologischer Reihenfolge vom ersten Ereignis eines Typs bis zur Gegenwart zu rekonstruieren. Für jede Episode sollten sie verschiedene inhaltliche Detailinformationen und zuletzt den Anfangs- und Endmonat angeben.

Abbildung 3.8: Ein Überblick über die Erhebungen der Deutschen Lebensverlaufsstudie (DLV)



Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin.

Zuerst wurden dabei Angaben zur Herkunftsfamilie erhoben, gefolgt von Wohn-, Schul-, Ausbildungs- und Erwerbsgeschichte sowie Informationen zu einigen weiteren Bereichen. Manche Bereiche wurden nur für bestimmte Gruppen der Kohorten erhoben (etwa die Zugehörigkeit zu staatlichen Organisationen der DDR für die ostdeutschen Teilnehmer). Am Ende lagen für jeden Teilnehmer mehrere lückenlose Ereignissequenzen innerhalb der erhobenen Bereiche vor.

Die Rekonstruktion der lebensgeschichtlichen Ereignissequenzen wurde in den ersten Befragungen durch Papierfragebögen in persönlichen Interviews bei den Befragten zu Hause vorgenommen; später telefonisch unter Einsatz eines rechnergestützten Systems zur Datenauf-

Abbildung 3.9: Erhebung der Erwerbsgeschichte in der LV-West

- 
- 1 Welchen Beruf haben Sie [in Ihrer ersten Stelle /Enddatum aus 3/ aus dem letzten Spell] ausgeübt?**
- 2 Welche berufliche Stellung hatten Sie in Ihrer damaligen Tätigkeit?**  
*INT: Vorgaben vorlesen*  
 1: Arbeiter(in) (z.B. in Produktion, Handwerk, Bau, Landwirtschaft)  
 2: Angestellte(r) (auch Angehörige(r) der Polizei, Armee)  
 3: Beamte  
 4: Genossenschaftsbauer bzw. selbständige(r) Landwirt(in)  
 5: Selbständig(e) in einem akademisch freien Beruf  
 6: Selbständig(e) (z.B. in Handel, Industrie, Dienstleistung oder Handwerk)  
 7: Mithelfende(r) Familienangehörige(r)  
 8: Heimarbeiter(in)  
 - 8: weiß nicht
- 3 Von wann bis wann waren Sie in dieser beruflichen Stellung tätig?**
- 4 Wie viel haben Sie im [Startdatum aus 3], zu Beginn der Tätigkeit, monatlich netto – also nach Abzügen – verdient?**
- 5 Was haben Sie direkt im Anschluß an diese Tätigkeit nach [Enddatum aus Q202] gemacht? Haben Sie ...**  
 1: die Erwerbstätigkeit unmittelbar danach fortgesetzt  
 2: Erwerbstätigkeit auch nur für kurze Zeit unterbrochen  
 3: oder ganz beendet?
- 5a Wenn Antwort 1: Haben Sie [Enddatum aus 3] dann in diesem Betrieb ...**  
 1: Ihre Tätigkeit gewechselt  
 2: Ihre berufliche Stellung verändert  
 3: oder haben Sie den Betrieb gewechselt?
- 5b Wenn Antwort 2 oder 3: Was haben Sie ab [Enddatum letzter Spell] genau gemacht? Waren Sie ...**  
 1: arbeitslos – arbeitssuchend  
 2: in Ausbildung – Umschulung – Studium  
 3: Hausfrau/-mann  
 4: im Erziehungsurlaub  
 5: im Vorruhestand – Ruhestand  
 6: Sonstiges
- 6 Von wann bis wann waren Sie [Kategorietext aus 5b], bis sich bei Ihnen eine Veränderung, z.B. durch eine neue Stelle oder durch eine Umschulung etc., ergeben hat?**
- 7 Haben Sie direkt [Enddatum aus 6] dann Ihre Erwerbstätigkeit wieder aufgenommen oder waren Sie weiterhin nicht erwerbstätig?**  
 1: Ja, wieder erwerbstätig (*zurück zur Erwerbsgeschichte*)  
 2: Nein, nicht mehr erwerbstätig (*Ende der Erwerbsgeschichte*)
- 

Ein Ausschnitt aus der Befragungsstruktur zur Erhebung der Erwerbsgeschichte in der letzten Erhebungswelle der LV an den westdeutschen Geburtskohorten 1964 und 1971.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin.

zeichnung (CATI). Die ersten Erhebungen in Ostdeutschland erfolgten wiederum durch persönliche Interviews, diesmal aber unter der Verwendung transportabler Computer zur Dateneingabe (CAPI). Die Erhebungsinstrumente wurden sehr stark strukturiert und enthielten sowohl die Fragen, die den Befragten gestellt wurden, als auch genaue Anweisungen, was wo wie aufnotiert oder verkodet werden sollte. Abbildung 3.9 zeigt einen Ausschnitt aus der Fragenstruktur der telefonischen Erhebung der Berufsgeschichte der LV-West. Die Interviews erreichten teilweise eine erhebliche Länge – einzelne Interviews der ältesten Geburtskohorten dauerten über fünf Stunden (Brückner, 1993), die meisten aber um die zwei Stunden.

Das Interview und die Erhebungsinstrumente wurden in Pilotstudien und Vortests so gestaltet, dass die Rekonstruktionsarbeit der Befragten unterstützt wurde (Mayer & Brückner, 1989; Papastefanou, 1980; Tölke, 1980). Auf einen Kalender als Datenaufzeichnungsinstrument wurde nach der Erprobungsphase allerdings verzichtet, da dessen Überlegenheit nicht eindeutig nachgewiesen werden konnte (Papastefanou, 1997; Tölke, 1980). In der Erhebung der LV werden folgende hilfreiche Erhebungsbedingungen umgesetzt:

*(a) Erhebung bereichsspezifischer Episodensequenzen*

Autobiografische Information wird in der LV erhoben, indem innerhalb eines Bereichs Episoden gebildet und durch inhaltliche und zeitliche Detailinformationen näher beschrieben werden. Dies entspricht sehr gut der Repräsentationsstruktur des autobiografischen Gedächtnisses, innerhalb dessen allgemeinere und spezifischere Episoden untereinander verbunden und zusammen mit Detailsindrücken repräsentiert sind. In der LV werden dabei Episoden erfragt, die den allgemeinen und zeitlich ausgedehnten Lebensabschnitten (im Sinne von Conway & Pleydell-Pearce, 2000) entsprechen. Diese sind entlang thematischer Achsen wie „Erwerbsleben“ oder „Familienleben“ organisiert, die den in der LV abgefragten Bereichen gut entsprechen. Der Einstieg über Lebensbereiche und -abschnitte hat sich als der effektivste Ansatzpunkt für die Rekonstruktion spezifischerer autobiografischer Information erwiesen (Conway & Bekerian, 1987).

*(b) Reihenfolge des Abrufs*

Da es sich um die Erhebung von Ereignissequenzen handelt, die sich entlang einer natürlichen Zeitachse entwickeln und deshalb in der Erinnerung auch als solche repräsentiert sind, ist die vorwärtsgerichtete chronologische Ordnung innerhalb der Lebensbereiche eine hervorragende Art, individuelle narrative Rekonstruktionspfade und *order codes* nutzbar zu machen.

In der LV wird jede Ereignissequenz vollständig erfasst, das heißt vom ersten Ereignis bis zur Gegenwart. Dies kann unter Umständen das ganze Leben des Befragten umfassen und somit eine erhebliche Anforderung an die Rekonstruktion darstellen. Auf der anderen Seite hat es Vorteile, vollständige und zum Teil lebenslange Sequenzen zu erheben: Wenn Sequenzen unvollständig erhoben werden, kann es sein, dass Ereignisse, die eigentlich korrekt erinnert werden (könnten), nicht angegeben werden, weil der Befragte sie versehentlich aus dem Zeitraum „hinausdatiert“ hat (Bradburn, Rips, & Shevell, 1987; Loftus & Marburger, 1983; Neter & Wakesberg, 1964; Rubin & Baddeley, 1989). Diese Verluste werden durch die vollständige Sequenzrekonstruktion verhindert.

*(c) Individuelle Kontextualisierung*

In der LV wird eine große Fülle autobiografischer Informationen aus vielen Lebensbereichen erhoben. Dadurch werden dem Befragten individuelle Kontexte, Referenzereignisse und Querverweise zugänglich gemacht. Dies erhöht die Menge der verfügbaren Abruf-Hinweise und Grenzsteine und fördert die Nutzung paralleler Abrufwege. Insofern ist dieses Vorgehen sowohl dem Erinnern als auch dem Datieren förderlich. Dadurch, dass die Lebensbereiche getrennt abgefragt werden, entfällt leider die Möglichkeit, Ereignisse und Kontexte aus anderen Lebensbereichen – das heißt parallele Rekonstruktionspfade – in vollem Umfang zielgerichtet als Abrufhilfe zu nutzen.

Wo die Datenaufzeichnung rechnergestützt erfolgte, gab es im System häufig die Möglichkeit, Fragen unter Bezugnahme auf zuvor gemachte Angaben zu formulieren, die für den Interviewer eingeblendet wurden („Sie sagten, bis November 1990 haben Sie als – *Antwort auf frühere Frage einblenden* – gearbeitet. Was haben sie danach gemacht?“). Solche sehr individuellen und spezifischen Abruf-Hinweise unterstützen die Rekonstruktion optimal und reduzieren außerdem die Wahrscheinlichkeit, dass ein Ereignis mit einem anderen verwechselt wird.

*(d) Grenzsteine*

Referenzereignisse zur Periodenbegrenzung fanden in der LV an einer Stelle Verwendung: Mit denjenigen Teilnehmern der ostdeutschen Teilstudien, die einer Rekontakierung zugestimmt hatten, wurde in den Jahren 1996 und 1997 eine erneute Befragung durchgeführt, um die Veränderungen seit der ersten Befragung 1991/92 nachzuzeichnen. Bei der Erhebung von Panelwellen hat sich die Technik des *bounded recall* als hilfreich erwiesen: Hier werden die Befragten noch einmal an ihre Angaben aus der vorherigen Panelwelle erinnert, um den Zeitraum, über den sie berichten sollen, eindeutig nach hinten abzugrenzen (Neter & Wakesberg, 1964; Sudman, Finn, & Lannom, 1984). In der LV wurde stattdessen für alle Teilnehmer gleichermaßen der Dezember des Jahres 1989 und das Ereignis des Mauerfalls gewählt. Da dieser Augenblick einen sehr markanten und folgenreichen Einschnitt in die Lebensumstände der ostdeutschen Befragten darstellt, sollte es nicht schwer sein, Ereignisse zeitlich vor oder nach dem Mauerfall einzuordnen.

*(e) Erhebung von Zeitinformation*

Kalenderdaten sind schwer zu erinnern und werden in der überwältigenden Mehrheit der Fälle inferenziell rekonstruiert. In der LV wurden deshalb die kalendarischen Daten jedes genannten Ereignisses als Allerletztes, nach allen anderen inhaltlichen Detailinformationen, erhoben. Dadurch standen den Befragten bereits die größtmögliche Fülle an temporal auswertbaren Informationen zur Zeitrekonstruktion zur Verfügung. Außerdem konnten die Befragten bei zu großer Unsicherheit an Stelle eines genauen Monats eine Jahreszeit angeben. Dies kommt den rekonstruktiven, auf viele komplementär und flexibel einsetzbaren Strategien beruhenden Datierungsmechanismen sehr entgegen und sollte gute Ergebnisse erzielen.

*(f) Absichernde Überprüfungsmöglichkeiten*

Plausible, innerhalb und zwischen Bereichen konsistente Lebensverläufe zu rekonstruieren und diese Konsistenz kontinuierlich zu überwachen, ist sowohl für den Befragten als auch für



den Interviewer eine komplexe Aufgabe. Fehler können sich einschleichen, die nur auf Flüchtigkeit beruhen und durch eine kleine Nachfrage sofort zu beheben wären. In den Befragungen mit Papierbögen hatten die Interviewer den Freiraum, etwa in späteren Teilen des Interviews, wenn eine Inkonsistenz auffiel, zu früheren Angaben zurückzugehen und Fehler zu beheben. Oft waren sie damit aber überfordert. In den rechnergestützten Befragungen wurden deshalb zahlreiche Konsistenz- und Plausibilitätsprüfungen schon während der Dateneingabe am Rechner durchgeführt, indem der vom Interviewer eingegebene Wert vom Programm mit bereits eingegebenen Daten oder zuvor festgesetzten Plausibilitätswerten und -wertebereichen verglichen wurde. Bei Inkonsistenzen werden dem Interviewer Fehlermeldungen eingespielt, etwa „Heiratsdatum Mutter muß nach Geburtsdatum Mutter liegen“ oder „Alter bei Geburt des ersten Kindes wirklich 45?“

*(g) Datenedition*

Um für die Weiterverarbeitung lückenlose, konsistente und plausible Rohdaten zu sichern, führte das Projekt vor der Dateneingabe eine umfassende Datenedition durch (Brückner, 1995; Hillmert, 2002; Wagner & Visser, 2004). In deren Rahmen wurden die Ereignissequenzen auf Lückenlosigkeit, Plausibilität und Widerspruchsfreiheit der Angaben innerhalb von und zwischen Lebensbereichen überprüft. Inkonsistente oder unplausible Angaben wurden soweit wie möglich aufgrund von Angaben aus anderen Teilen des Interviews rekonstruiert. Wenn dies nicht möglich war, wurde mit dem Interviewten telefonisch, in Einzelfällen auch schriftlich, eine klärende Nachbefragung durchgeführt, was in etwa 15 bis 20 Prozent der Fälle notwendig wurde. Besonderer Wert wurde auf die Transparenz der Dateneingriffe gelegt. Änderungen wurden im Datensatz immer kenntlich und anhand der Protokolle nachvollziehbar gemacht. „Editing, then is neither more or less than an attempt to compensate for the shortcomings of the [interviewing] process by reordering and reconsidering data.“ (Brückner, 1995)

Reliabilitäts- und Validierungsstudien können dabei mit den edierten oder den unedierten Daten arbeiten. Da ich mich für die Ergebnisse der Erinnerungsarbeit der Befragten in der Befragungssituation unter den gegebenen realistischen Bedingungen interessiere, liegt die Arbeit mit den unedierten Datensätzen näher. Diese enthalten die autobiografischen Rekonstruktionen des Befragten so, wie dieser sie unter den Bedingungen der Befragung vorgenommen hat. Einschränkend muss man anmerken, dass in Einzelfällen auch die Angaben im edierten Datensatz diesem abgelieferten Erinnerungsprodukt näher kommen als die unedierten: So kann es durchaus vorkommen, dass Befragte bestimmte Angaben zwar machten, diese aber vom Interviewer nicht aufgezeichnet wurden und erst durch das Abhören der Tonbandmitschnitte der Sitzung (soweit vorhanden) nachgetragen werden können.

*(h) Weitere Punkte*

Das Konzept des Lebensverlaufs als Sequenz von Zuständen, die sich innerhalb eines Bereichs gegenseitig ausschließen und den Zeitraum lückenlos beschreiben können, stellt eine idealisierte Version eines Lebenslaufs dar, dem sich die tatsächlichen Lebensverläufe nicht immer gänzlich fügen. Eine Person, die beispielsweise einer betrieblichen Ausbildung nachgeht, steht vor der Wahl, diesen Zeitraum entweder als Ausbildung zu bezeichnen und damit den Erwerbscharakter des Zeitraums zu ignorieren oder ihn unter Erwerbstätigkeit zu führen und den Ausbildungscharakter zu ignorieren. Ein Befragter, der zugleich Kinder aufzieht und den

Haushalt versorgt und eine freiberufliche Tätigkeit ausübt, muss sich entscheiden, ob er sich als Hausmann oder als Erwerbstätiger einstuft. Um die Rigidität des Sequenzkonzepts abzuschwächen, wurden deshalb innerhalb eines Bereichs oft zwei verschiedene Verlaufspfade unterschieden. Der eine repräsentiert stärker institutionell eingebundene, konventionellere Ereignisse, die eindeutig abgrenzbar direkt aufeinander folgen; der andere informellere, weniger institutionell geregelte und unkonventionellere Verläufe, bei denen Ereignisse sich auch überlappen können. So wurde etwa sowohl nach leiblichen Eltern als auch nach Konstellationen von Stief- oder Zieheltern gefragt und sowohl nach Erwerbstätigkeiten als auch nach Nebenerwerbstätigkeiten. Außerdem wurde in der letzten Erhebung der Kohorten von 1964 und 1971 die Gleichzeitigkeit von Erwerbstätigkeiten und Ausbildungsphasen zugelassen, da dies in den jüngeren Kohorten zunehmend häufiger vorkommt.

In vielerlei Hinsicht stehen die Anforderungen, die die LV an die Rekonstruktionsarbeit der Befragten stellt, im Einklang mit den Rekonstruktionsmechanismen des autobiografischen Gedächtnisses. Sie erlauben also den Stärken der Erinnerung, sich zu entfalten, und bieten ihren Schwächen spezifische und individuelle Unterstützung. Deshalb lässt sich vermuten, dass die erhobenen Daten trotz der hohen Anforderungen an die Befragten von hoher Qualität sind (Brückner, 1993).

Die LV bietet zwei hervorragende Gelegenheiten, die Fehler beim Rekonstruieren systematisch und differenziert zu beschreiben und im Rahmen spezifischer Annahmen über die ablaufenden Rekonstruktionsprozesse als Vereinfachung und Konventionalisierung zu erklären: eine Wiederholungsbefragung und ein Registervergleich.

In der ostdeutschen Teilstudie der LV wurde ein Teil der Befragten zu Teilen ihrer Lebensgeschichte nach vier bis fünf Jahren erneut in identischer Form befragt. Durch einen Vergleich lässt sich also die Reliabilität der Angaben abschätzen. Ich werde in Kapitel 4 die Übereinstimmung der Angaben über die Kinder untersuchen, und zwar das Auslassen von Kindern und Fehler beim Datieren von Geburt und Auszug der Kinder. Geburten und Auszüge stellen an die Datierung sehr unterschiedliche Anforderungen, sodass ich unterschiedlich verbreitete, große und anders gerichtete Fehler erwarte. Außerdem werde ich betrachten, ob die Datierungen für länger zurückliegende Auszüge wirklich schlechter ausfallen als für kürzer zurückliegende Auszüge, und werde die Form des Zusammenhangs zwischen retrospektiver Distanz und Fehler beim Datieren des Auszugs näher untersuchen.

In Kapitel 5 werde ich die Übereinstimmung der Angaben zu den Erwerbsverläufen Anfang der 1990er Jahre betrachten. Durch die rasante und radikale Transformation des Arbeitsmarkts und des Beschäftigungssystems in Ostdeutschland liegen für diesen Zeitraum spezielle Bedingungen für die Rekonstruktion von Erwerbsbiografien vor. Insbesondere erwarte ich, dass sich die Angaben auf eine Art und Weise zwischen den beiden Befragungen unterscheiden, die nicht (nur) das zunehmende Vergessen und Umdeuten widerspiegeln, sondern auch die Umwälzungen in den Erwerbsaktivitäten und deren Wahrnehmung und Einordnung. Ich werde vor allem die Reliabilität der zeitlichen Information berücksichtigen, die in den Daten enthalten ist, nämlich die Dauer von Erwerbsepisoden und die Häufigkeit und Datierung von Übergängen. Neben den Ursachen und Folgen der Abweichungen werde ich außerdem untersuchen, welche Verbesserungen durch die Datenedition erzielt wurden.

Für zwei Kohorten aus Westdeutschland lassen sich die Angaben auf Personenebene mit den Daten der Bundesanstalt für Arbeit verknüpfen. Dieser Vergleich erlaubt eine Abschätzung der Validität in einem vollständigen Registerdatenvergleich. In Kapitel 6 werde ich

mich eingehend mit der Übereinstimmung der Angaben zu Arbeitslosigkeit mit den Registereinträgen beschäftigten, deren aussagekräftige und vergleichbare Erhebung in Befragungen und Verwaltungsdaten ein vieldiskutiertes Thema in der Arbeitsmarktforschung ist (Elias, 1997; Norwood & Tanur, 1994). Ich werde zum einen untersuchen, wie gut amtlich registrierte Perioden des Leistungsbezugs erinnert werden, und zum anderen, unter welchen Umständen Zeiträume ohne Leistungsbezug dennoch als Arbeitslosigkeiten angegeben werden. Außerdem werde ich betrachten, ob die Übereinstimmung für weiter zurückliegende Episoden schlechter wird und welche Form der Zusammenhang zwischen retrospektiver Distanz und Übereinstimmung hat.

Ich werde für jeden Vergleich ein partielles Modell des Datengenerierungsprozesses aufstellen, in dem für die konkrete Anforderung die kognitiven Schritte, die der Befragte ausführen muss, herausgearbeitet werden. Aus diesem Modell werde ich Annahmen über das Auftreten und die Art von Datenfehlern ableiten. Ich werde den Datenfehler umfassend und differenziert hinsichtlich aller vier Aspekte (Verbreitung, Größe, Richtung und Abhängigkeit) beschreiben, und jeweils unter Bezug auf das Datengenerierungsmodell als Folge von Erinnerungsprozessen erklären.

## Kapitel 4

### „Haben Sie Kinder?“ Die Reliabilität der Angaben zu Kindern in der LV-Ost

#### 4.1 Fragestellung und Vorgehen

Die Information darüber, ob und wenn ja wie viele Kinder eine Person hat und wie sich das Zusammenleben mit diesen Kindern im Zeitverlauf und in Bezug auf die Verläufe in anderen Lebensbereichen gestaltet, ist für das Verständnis von Lebensverlaufsprozessen in vielerlei Hinsicht zentral. Wichtige Lebensentscheidungen werden anders gefällt, wenn Kinder geplant oder vorhanden sind, sodass die ökonomische und soziale Situation von Müttern und Vätern in hohem Maße vom Vorhandensein und der Anzahl von Kindern geprägt werden. Etwa führt die Geburt von Kindern sehr oft dazu, dass ihre Mütter ihre Erwerbstätigkeit unterbrechen und dadurch Nachteile in der Einkommens- und Karriereentwicklung erfahren (Huinink, 1995; Lauterbach, 1994). Auch das Scheidungsrisiko von Eltern sinkt im Vergleich zu kinderlosen Paaren, nimmt aber ab einer gewissen Kinderzahl wieder zu, bis Eltern und kinderlose Paare wieder gleichauf liegen (Heaton, 1990; Thornton, 1977). Auch der Zeitpunkt, an dem sie in die Familie eintreten bzw. sie wieder verlassen, spielt eine Rolle für die elterlichen Berufs- und Beziehungsverläufe – etwa indem sowohl die Erwerbstätigkeit der Mütter als auch das Scheidungsrisiko der Eltern mit dem Alter der Kinder wieder ansteigt (Ott, 1991; Wagner, 1996a). Der Auszug von Kindern wird auch als das Phänomen der verzögerten Haushaltsgründung von Kindern, die „das Nest nicht verlassen“, untersucht (Hullen, 1995; Lauterbach & Lüscher, 1999; Mayer & Wagner, 1986; Papastefanou, 1997). Im Zuge des gesellschaftlichen Wandels und der damit einhergehenden Transformation der „klassischen“ Familie rücken auch zunehmend nichtleibliche Kinder ins Zentrum der Aufmerksamkeit – etwa Kinder des Partners oder eigene Kinder, die im Haushalt des getrennten Partners leben – und deren Einfluss auf das Scheidungsrisiko der sie aufziehenden Paare (Wagner, 1996a). Auch hier spielt der Zeitverlauf des Zusammenlebens, also der Zeitpunkt, an dem nichtleibliche Kinder in den Haushalt eintreten bzw. ihn wieder verlassen, eine wichtige Rolle (Morgan & Rindfuss, 1985). Deshalb ist es von großer Bedeutung, dass alle Kinder und die Zeiträume, in denen sie mit den Befragten zusammengelebt haben, korrekt erfasst werden.

In diesem Kapitel betrachte ich die Angaben zur Erwerbsbiografie aus der Wiederholungsbefragung der *Deutschen Lebensverlaufstudie* (LV) in Ostdeutschland. Die beiden Befragungen wurden in einem Abstand von vier bis fünf Jahren durchgeführt. Ich werde mich beim Vergleich auf folgende Fragen konzentrieren:

- Werden mit längerer retrospektiver Distanz Kinder ganz unterschlagen?
- Wie zeitstabil werden die Daten von Geburtstag und Auszug der Kinder rekonstruiert?
- Weichen die Daten länger zurückliegender Auszüge stärker voneinander ab?

Zunächst werde ich einen Überblick über die Befundlage zur Qualität kindbezogener Befragungsdaten geben. Danach werde ich beschreiben, wie kindbezogene Information in der LV

erhoben wird und auf dieser Grundlage im Sinne eines partiellen Datengenerierungsmodells die kognitiven Schritte herausarbeiten, die die Befragten vollführen müssen. Aus diesem werde ich Erwartungen über die Abweichungen zwischen den Befragungen ableiten. In den Abschnitten 4.3.1 und 4.3.2 werde ich den Fehler als Irrtum differenziert im Hinblick auf *Verbreitung*, *Ausmaß* und *Richtung* beschreiben und durch logistische Regression untersuchen, ob das Risiko, Kinder ganz zu unterschlagen, sich zwischen Gruppen von Befragten und zwischen Gruppen von Kindern unterscheidet. In Abschnitt 4.3.3 werde ich für die Datierung der Auszüge der Kinder untersuchen, ob länger zurückliegende Auszüge häufiger fehldatiert werden, ob das Ausmaß der Fehldatierungen zunimmt und sich die Richtung der Fehldatierungen verändert. Da die zu erinnernde Information ebenfalls eine Zeitinformation ist, kann man dies auch als *Abhängigkeit* der Fehleraspekte vom wahren Sachverhalt betrachten. In Abschnitt 4.4 werde ich die Befunde zusammenfassen und diskutieren.

## 4.2 Retrospektivbefragungen zu Kindern: Befunde und Annahmen

### 4.2.1 Befundlage

Studien, die die Angaben der Befragten zu ihren Kindern in Wiederholungsbefragungen oder durch Registerdatenvergleiche überprüfen, zeigen übereinstimmend, dass Kinder äußerst selten vergessen werden. Auriat (1992) fand keinen solchen Fall bei Vergleich der Angaben der Befragten der „Enquête 3B.bis“ mit dem belgischen Bevölkerungsregister, und Papastefanou (1980) stellte bei einem Vergleich der zehn Jahre auseinander liegenden Angaben in der „Konstanzer Mobilitätsstudie“ fest, dass 94 Prozent der (männlichen) Befragten die gleiche Anzahl Kinder angaben.

Beim Vergleich der Angaben der Befragten der „Enquête 3B.bis“ mit dem belgischen Bevölkerungsregister fanden Poulain, Riandey und Firdion (1991), dass die Geburtsdaten der Kinder in 98 Prozent der Fälle korrekt angegeben wurden, wobei Männer sich ein wenig häufiger irren und nur in 94 Prozent der Fälle korrekte Angaben machten. Meist wurde der Monat korrekt und das Jahr um ein, sehr selten um zwei Jahre abweichend angegeben. Solche „Skaleneffekte“ entstehen, weil Monat und Jahr unabhängig voneinander erinnert werden und weil die Rekonstruktion des Monats leichter fällt als die des Jahres (siehe Abschnitt 2.5.3). Auch hier waren Frauen ein wenig besser als ihre Ehepartner.

Anders sah es aus für die Auszugsdaten: Von den in den Registern vermerkten Auszügen der Kinder wurden von sowohl Männern als auch Frauen nur etwa 90 Prozent überhaupt angegeben, und hier waren die Angaben von Männern geringfügig besser. Von den angegebenen Auszügen wurden 56 Prozent von Männern und 48 Prozent von Frauen fehldatiert. Die Größe der Abweichungen ging aus der Publikation leider nicht hervor, aber Abweichungen von über einem Jahr traten ebenso häufig auf wie Abweichungen von unter einem Jahr. Skaleneffekte fanden sich keine, das heißt, dass sich die Abweichungen nicht im Jahresabstand zum Registerdatum häuften. Außerdem fanden die Autoren eine deutliche Neigung zum Vordatieren der Auszüge bei beiden Geschlechtern, die für Frauen aber noch stärker war: Diese gaben im Falle einer abweichenden Angabe doppelt so häufig ein zu frühes Datum an.

Die Fehler beim Datieren von Auszügen sind also nicht nur deutlich häufiger und größer, sie verteilen sich auch anders und treten gerichtet auf. Dies deutet darauf hin, dass es sich bei

der Datierung von Auszügen um einen anders gelagerten Fall als bei der Datierung von Geburtstagen handelt und dass andere Mechanismen an der Fehlerproduktion beteiligt sind.

#### 4.2.2 Datengenerierungsmodelle und Annahmen für die LV-Ost

##### *Datengenerierung*

Nach den umfangreichen und erfolgreichen Erhebungen der LV mit westdeutschen Teilnehmern wurde unter dem Motto „Lebensverläufe und historischer Wandel in der ehemaligen DDR“ nach dem Fall der Mauer die Erhebung auf das Gebiet der ehemaligen DDR ausgedehnt. Mit einem an die institutionellen Gegebenheiten angepassten Fragenkatalog wurden in der Haupterhebung zwischen September 1991 und August 1992 insgesamt 2.331 persönliche Befragungen<sup>1</sup> mit Mitgliedern der Geburtsjahrgänge 1929–31, 1939–41, 1951–53 und 1959–61 durchgeführt. Mit denjenigen Teilnehmern ( $N = 1394$ ), die einer Rekontaktilierung

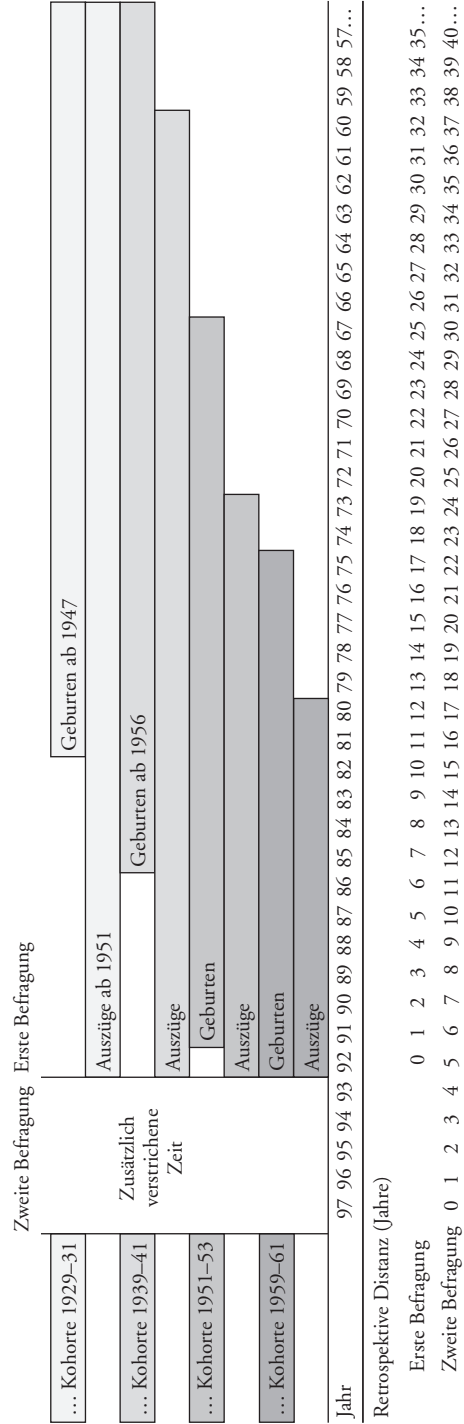
Abbildung 4.1: Erhebung von Angaben zu Kindern in der LV-Ost

- 
- 1 **Haben Sie Kinder? Ich meine sowohl eigene Kinder als auch Pflege-/Adoptivkinder oder Kinder des/der Partner(in), die bei Ihnen aufwachsen bzw. aufgewachsen sind. Bitte denken Sie dabei auch an solche, die bereits verstorben sind.**  
 1: Ja  
 2: Nein
- Wenn die ZP hier mit nein antwortet, in der Wohngeschichte jedoch angegeben hat, mit einem Kind zusammenzuwohnen, so wird der Interviewer angehalten, noch einmal nachzufragen und die Antwort ggf. zu korrigieren*
- 2 **Wie viele Kinder haben Sie?** \_\_\_\_\_
- 3 **Wenn nur ein Kind:** Nun möchte ich Ihnen einige Fragen zu Ihrem Kind stellen. Sagen Sie mir bitte der Einfachheit halber den Vornamen und nennen Sie mir Geschlecht und Geburtsdatum des Kindes.  
**Bei mehreren Kindern:** Nun möchte ich Ihnen einige Fragen zu Ihren Kindern stellen. Beginnen Sie mit Ihrem ältesten Kind und sagen Sie mir auch bitte gleich, wann dieses Kind geboren wurde. Der Einfachheit halber bitte ich Sie, mir auch die Vornamen zu nennen.
- 4 **Wie heißt das erste (nächste) Kind mit Vornamen?** \_\_\_\_\_  
**Welches Geschlecht?** \_\_\_ 1 -> männlich 2 -> weiblich  
**Und welches Geburtsdatum?** Tag \_\_\_ Monat \_\_\_ Jahr 19 \_\_\_
- 5 **Ist [Name Kind] Ihr ...**  
 1: leibliches Kind  
 2: Kind des Partners  
 3: Adoptivkind  
 4: Pflegekind
- 6 **Lebt [Name Kind] noch?**  
 1: Ja  
 2: Nein
- 7 **Wann ist [Name Kind] ausgezogen? Nennen Sie mir bitte das Auszugsjahr.**  
 Jahr des Auszugs \_\_\_\_\_
- 

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin.

<sup>1</sup> Die Durchführung telefonischer Interviews war in diesem Fall nicht möglich, da im Erhebungszeitraum (September 1991 bis Oktober 1992) die Telefondichte in den neuen Ländern nicht hoch genug war, um eine für die Gesamtbevölkerung hinreichend repräsentative Stichprobe zu ziehen.

Abbildung 4.2: Zeiträume, aus denen Angaben zu Geburts- und Auszugsdaten der Kinder vorliegen



Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

zugestimmt hatten, wurde in den Jahren 1996 und 1997 eine erneute Befragung durchgeführt, um die Veränderungen seit der ersten Befragung 1991/92 nachzuzeichnen. Dabei wurden die Angaben zu Kindern noch einmal komplett abgefragt; der Wortlaut der Fragen ist in Abbildung 4.1 dargestellt. Die Frage nach den Kindern betont dabei explizit, dass auch bereits gestorbene Kinder, Adoptiv- und Pflegekinder sowie Kinder des Partners, sofern sie mit dem Befragten zusammenwohnen oder gewohnt haben, angegeben werden sollen. Von jedem Kind wird zuerst der Vorname erfragt und dann alle weiteren Fragen unter Nennung dieses Namens gestellt; so kann es weniger leicht zu einer Verwechslung der Kinder kommen. Abbildung 4.2 zeigt den zweifach abgedeckten Zeitraum und wie sich für die vier befragten Kohorten Geburten und Auszüge verteilen.

Während die Daten für die inhaltlichen Analysen einer intensiven und sorgfältigen Edition zur Fehlerreduktion unterzogen wurden (Brückner, 1995; Brückner & Mayer, 1995; Wagner & Visser, 2004), verwende ich hier die uneditierten Datensätze, in denen noch alle (Gedächtnis-)Fehler in vollem Umfang enthalten sind. Diese stellen originäre Produkte der Erinnerungsarbeit der Befragten in der vom Projekt konzipierten und gestalteten Abrufsituation dar. Darüber hinaus wurden in der Edition der zweiten Befragung teilweise die Angaben aus der ersten Befragung herangezogen, um möglicherweise falsche Angaben zu identifizieren und entweder anzugleichen oder in der Nachrecherche aufzuklären. Ein Vergleich der edierten Datensätze würde also ein unrealistisch hohes Maß an Übereinstimmung ergeben.

#### *Kinder: Wie werden sie erinnert?*

In Abbildung 4.3 ist dargestellt, welche kognitiven Schritte der Befragte vollführen muss, um sich an seine Kinder zu erinnern: den Begriff „Kinder“ interpretieren und mit diesem Begriff als Abruf-Hinweisreiz in ihrem autobiografischen Gedächtnisbestand nach solchen Kindern suchen. Da Kinder die Lebenssituation und sozialen Rollen ihrer Eltern nachhaltig und massiv verändern, nehme ich an, dass Kinder nicht wirklich vergessen werden. Allerdings ist es denkbar, dass bei der Interpretation der Kategorie „Kinder“ Abweichungen auftreten und selektiv ganz bestimmte Kinder manchmal nicht in die Definition aufgenommen werden, nämlich bereits *gestorbene* und *nichtleibliche* Kinder (solche, die adoptiert, als Pflegekinder

Abbildung 4.3: Datengenerierungsprozess bei der Erhebung von Kindern, ihrer Geburts- und Auszugsdaten

Rekonstruktion der Kinder	
Schritt 1: Abruf-Hinweisreiz „Kinder“ interpretieren	
Schritt 2: Kinder erinnern	
Rekonstruktion des Geburtsdatums	
<i>entweder:</i>	
Schritt 1: Abruf-Hinweisreiz „Geburtstag des Kindes“ interpretieren	
Schritt 2: Geburtstag des Kindes rekonstruieren	
Schritt 3: Monat des Geburtstages rekonstruieren	
Schritt 4: Jahr des Geburtstages rekonstruieren	
<i>oder:</i>	
Schritt 1: Geburtsdatum direkt abrufen, wenn als über- lerntes Zeitetikett vorhanden	
Rekonstruktion des Auszugsdatums	
<i>entweder:</i>	
Schritt 1: Abruf-Hinweisreiz „Auszug des Kindes“ interpretieren	
Schritt 2: „Auszug des Kindes“ rekonstruieren	
Schritt 3: Jahr des Auszugs rekonstruieren	
<i>oder:</i>	
Schritt 1: Auszugsdatum direkt abrufen, wenn als über- lerntes Zeitetikett vorhanden	



aufgenommen oder vom Partner mit in die Beziehung gebracht wurden). Solche „Nichtstandardkinder“ sind in der Erinnerung der Befragten weniger stabil repräsentiert und werden auf den Hinweis „Kinder“ nicht mit der gleichen Zwangsläufigkeit aktiviert. Mit der verstrichenen Zeit dürfte der Fehler eine deutliche Richtung aufweisen: Kinder werden in den allermeisten Fällen zu späteren Zeitpunkten unterschlagen, nicht zusätzlich angegeben, da so die eigene Biografie konventioneller und kontinuierlicher erscheint.

Auf Seiten der erinnernden Personen ist anzunehmen, dass das Irrtumsrisiko für Männer höher ist als für Frauen, da der Lebensverlauf von Müttern noch massiver und einschneidender von Kindern geprägt wird als der von Vätern. Unterschiede zwischen den Kohorten erwarte ich keine, da auch die älteste Kohorte des Vergleichs noch nicht alt genug ist, um hier altersbedingte kognitive Beeinträchtigungen aufzuweisen. Allerdings nehme ich an, dass eine höhere Kinderzahl das Irrtumsrisiko erhöht.

#### *Geburtstage und Auszüge: Wie werden sie datiert?*

In Abbildung 4.3 ist dargestellt, welche kognitiven Schritte der Befragte vollführen muss, um Geburtstag und Auszug eines Kindes zu datieren. Kalenderdaten werden für einige besonders zentrale Grenzsteinereignisse durch wiederholtes Üben und Einprägen überlernt und können direkt abgerufen werden. Für alle anderen Ereignisse werden sie inferenziell und näherungsweise aus anderen Erinnerungen rekonstruiert, indem der Befragte zunächst das fragliche Ereignis und dann dessen Datum rekonstruiert. Abweichungen der Datumsangaben können also auf allen Schritten des Konstruktionsprozesses entstehen:

- wenn das zu datierende Ereignis abweichend interpretiert wird,
- wenn das zu datierende Ereignis abweichend rekonstruiert wird,
- wenn das (übereinstimmend rekonstruierte) Ereignis abweichend datiert wird (bei Geburtsdatum Monat und Jahr, bei Auszugsdatum nur das Jahr).

Für Geburtsdaten ist aus mehreren Gründen eine wesentlich zeitstabilere Rekonstruktion zu erwarten. Zum Ersten ist kaum vorstellbar, dass Abweichungen in der Rekonstruktion des Ereignisses „Geburt eines Kindes“ auftreten. Zum Zweiten markieren Kindsgeburten den Übertritt in einen neuen Lebensabschnitt der Befragten im familiären Bereich (Conway, 1996) und bilden Grenzsteine (Auriat, 1996). Sie werden also häufig mit direkt abrufbaren Zeitetiketten versehen, was eine hohe Rekonstruktionsstabilität bedeutet. Grenzsteine sind in den temporalen Strukturen des autobiografischen Gedächtnisses vielfältig verknüpft und reich an temporalen Abruf-Hinweisreizen. Außerdem werden Geburtstagsdaten, besonders der Monat, durch ihre alljährliche feierliche Begehung permanent geübt und eingepägt.

All dies ist für Auszugsdaten nicht oder in geringerem Maß der Fall. Schon beim ersten Schritt ist ein Auszug ein weniger eindeutig abgrenzbares, an einem einzigen Punkt in der Zeit lokalisiertes Ereignis; ein Kind kann beispielsweise seinen Lebensmittelpunkt in eine andere Stadt verlegen, aber bei den Eltern gemeldet bleiben, sei es aus Nachlässigkeit oder aus guten (z.B. versicherungsrechtlichen) Gründen. Auch können Kinder nach längeren Perioden mehr oder weniger vollständigen Auszugs wieder vorübergehend oder dauerhaft ins „Nest“ zurückkehren – etwa im Zusammenhang mit Wehr- oder Ersatzdienst, mit Ausbildungen oder beruflichen oder privaten Schwierigkeiten. Wenn dem Befragten mehrere Ereignisse als „Auszug“ zur Auswahl stehen, kann es dazu kommen, dass in den beiden Befragungen unterschiedliche Ereignisse datiert werden. Dies verursacht dann (selbst bei korrekter Datierung)

inkonsistente Datumsangaben. Auch beim Datieren selbst sind häufigere und größere Abweichungen zu erwarten, da sie weniger deutlich den Übertritt der Befragten in einen neuen Lebensabschnitt markieren. Auszüge der Kinder werden vermutlich seltener mit einem Zeitetikett versehen und nicht alljährlich begangen und eingeübt, müssen also meistens inferenziell rekonstruiert werden. Auch dies ist schwieriger als bei Geburtsdaten, denn sie sind weniger eng in die temporale Struktur des autobiografischen Gedächtnisses eingebunden. Alles dies bedeutet ein erhöhtes Irrtumsrisiko und größere Abweichungen beim Datieren von Auszügen.

Darüber hinaus könnte es sein, dass der Fehler bei Auszugsdaten gerichtet auftritt. Zwar fällt der Fehler beim Datieren selbst symmetrisch aus, wenn aber Fehler beim ersten Schritt, der Ereignisrekonstruktion, darauf hinauslaufen, dass systematisch spätere oder frühere Ereignisse als Auszug rekonstruiert werden, kann hier eine Fehlerrichtung eingeführt werden. So lässt sich die von Poulain u.a. (1991) gefundene Neigung der Befragten, im Vergleich zum Bevölkerungsregister den Auszug ihrer Kinder früher zu datieren, durch diesen Schritt der Datenkonstruktion eher erklären: Wenn Kinder sich nach dem physischen Verlassen des elterlichen Haushalts längere Zeit nicht behördlich ummelden, geben die Befragten ein zu frühes Datum an, wenn sie sich an diesen tatsächlichen Umzug erinnern<sup>2</sup>.

*Die retrospektive Distanz: Wie beeinflusst sie die Konsistenz?*

Während also Datumsangaben (und auch sonst alle Angaben) zum Auszug von Kindern generell instabil sind, bleibt die Frage, ob weiter zurückliegende Auszüge anders (häufiger, größer, anders

#### Abbildung 4.4: Hypothesen im Überblick

---

##### *Kinder:*

- werden nur selten inkonsistent erinnert
- werden stärker unterschlagen als hinzugefügt
- nichtleibliche Kinder und bereits gestorbene Kinder werden weniger konsistent erinnert und ebenfalls stärker unterschlagen als hinzugefügt
- Frauen erinnern sich konsistenter
- keine Kohortenunterschiede
- je mehr Kinder eine Person hat, desto eher wird eines unterschlagen

##### *Daten:*

- Geburtsdaten sehr selten inkonsistent; ein bis maximal zwei Jahre Differenz; Abweichungen symmetrisch; Skaleneffekte (Monat besser als Jahre)
- Auszugsdaten häufiger inkonsistent, häufig mehrere Jahre Differenz; Abweichungen symmetrisch
- Frauen datieren konsistenter

##### *Längere retrospektive Distanz:*

- Abweichungen häufiger
  - Abweichungen größer (lineare Zunahme)
  - frühere Auszüge eher nach später, spätere Auszüge eher nach früher fehldatiert
- 

<sup>2</sup> Die Autoren der oben genannten Studie meinen, dass diese Verteilung der Abweichungen „va à l'encontre d'une interpretation en terme d'erreur de mémoire“ (Poulain u.a., 1991, S. 82). Da sie zwar keine reinen *Datierungsfehler* darstellen, aber auf Stufe 1 des Rekonstruktionsprozesses zu Stande kommen, sehe ich keinen Grund, sie nicht als Erinnerungsfehler zu behandeln. Fehler in einem realistischen Kontext kommen häufig durch mehrere Mechanismen zu Stande und spiegeln nicht den Effekt eines isolierten Faktors wider.

gerichtet) abweichend definiert bzw. fehldatiert werden als weniger weit zurückliegende. Ich habe keinen Grund anzunehmen, dass bei Rekonstruktionsschritt 1 und 2 für länger zurückliegende Auszüge häufiger der Abruf-Hinweisreiz anders interpretiert oder andere Ereignisse rekonstruiert werden; da aber die Daten länger zurückliegender Auszüge schwieriger zur rekonstruieren sind, nehme ich an, dass Irrtumsrisiko und Fehlergröße über die Zeit zunehmen.

Rubin und Baddeley (1989) schlagen – wie in Abschnitt 2.5.3 beschrieben – vor, den Datierungsfehler in Analogie zu den Weber'schen Wahrnehmungsgesetzen als einen Fehler beim Einschätzen einer seit einem Ereignis verstrichenen Zeitspanne zu betrachten, der mit der zunehmenden retrospektiven Distanz deshalb *linear* zunimmt. Rubin und Baddeley sprechen von einer Zunahme des Datierungsfehlers um 0,15 Tage pro verstrichenem Tag, Linton (1975) von 0,22 Tagen und Thompson (1982) von einem Tag pro Woche. Dies bezieht sich auf Zeiträume bis zu sechs Jahren. Ich werde überprüfen, ob eine vergleichbare Linearfunktion auch für Auszüge der Kinder und retrospektive Distanzen bis zu 20 Jahre gilt.

Zur *Richtung* der Abweichungen nehme ich an, dass länger zurückliegende Ereignisse eher zu spät angegeben werden und kürzer zurückliegende Ereignisse eher zu früh (Thompson, Skowronski, Larsen, & Betz, 1996). Denn der Zeitraum, in den fehlerhafte Datierungen fallen können, ist durch die normativen Entwicklungslogiken eines Lebenslaufs begrenzt: Kinder ziehen unter einem bestimmten Alter (der Eltern bzw. der Kinder) einfach nicht aus und bis zu einem bestimmten Alter haben sie ausgezogen zu sein. Dadurch können die Ereignisse zu Beginn der Referenzperiode nicht beliebig viel früher datiert bzw. ein beliebig viel früheres Ereignis an seiner Stelle als Auszug definiert werden. Wenn ein Kind im Alter von 18 Jahren ausgezogen ist, wird sich vier Jahre *später* eher ein verwechselbares Ereignis finden als vier Jahre früher, und eine irrtümliche Fehldatierung um vier Jahre nach früher, der zufolge das Kind mit 15 Jahren ausgezogen wäre, würde so stark in Widerspruch mit der Normbiografie treten, dass sie eine Korrektur nach sich zöge. Deshalb treten Abweichungen früher Umzüge vorzugsweise in Richtung „später“ auf. Umgekehrt verhindert die normative Grenze, dass kürzer zurückliegende Auszüge noch näher an die Gegenwart heran verschoben werden bzw. dass ein noch kürzer zurückliegendes Datum rekonstruiert wird. Die Irrtümer laufen also darauf hinaus, dass Extreme nivelliert und die Angaben aus der längeren Rückschau konventioneller ausfallen. Da die Auszüge jünger ausgezogener Kinder auch länger zurückliegen, müssen die beiden Effekte des Alters des Kindes und der retrospektiven Distanz sorgfältig voneinander getrennt werden.

### 4.3 Auswertung

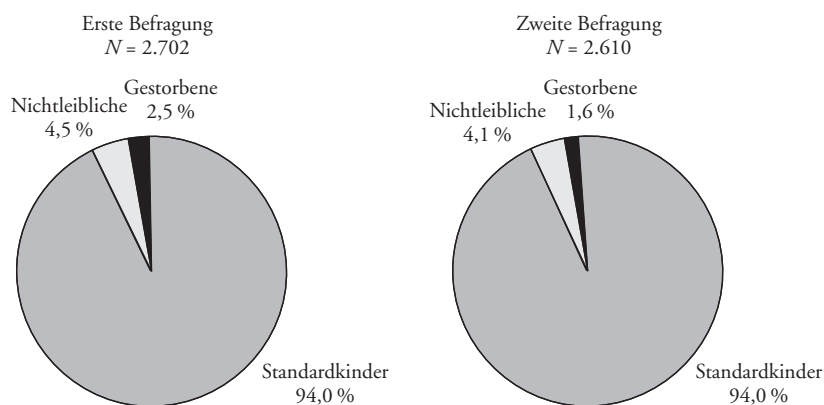
#### 4.3.1 Das Unterschlagen von Kindern

In der ersten Befragung werden von den 1.301 Befragten Angaben zu 2.702 Kindern gemacht, in der zweiten zu 2.610, was eine Differenz von 92 Kindern ergibt<sup>3</sup>. 89,2 Prozent der Teilnehmer geben beide Male dieselbe Anzahl von Kindern an, 8,6 Prozent geben unter-

<sup>3</sup> Die Angaben aus der zweiten Befragung wurden um diejenigen Kinder bereinigt, deren Geburtsdatum erst nach dem ersten Interview lag. Befragte, die in beiden Befragungen übereinstimmend angegeben hatten, kinderlos zu sein, wurden in dieser Analyse nicht berücksichtigt.

schiedliche Anzahlen von Kindern an, und 2,6 Prozent werden in einer der Befragungen als kinderlos geführt. Abbildung 4.5 zeigt die Verteilung der Kindschaftsverhältnisse und des Vitalstatus in den beiden Befragungen. Die überwältigende Mehrheit der Kinder in beiden Interviews ist leiblich und am Leben. Allerdings ist in der zweiten Befragung der Anteil der nichtleiblichen Kinder leicht und der Anteil der gestorbenen Kinder deutlich verringert. Um auf der Individualebene zu ermitteln, welche Kinder von wem vergessen werden, habe ich die Kinder aus den beiden Befragungen einander paarweise auf der Grundlage von Befragten-ID und Vornamen des Kindes zugeordnet, wobei ich bezüglich Rechtschreibung und Doppelnamen großzügig war („Klaus“ wurde als dasselbe Kind wie „Claus-Dieter“ gewertet, „Rayko“ als dasselbe wie „Heiko“, wenn sie jeweils vom selben Befragten angegeben worden waren). Für insgesamt 2.559 Kinder konnte so ein „Zwilling“ im anderen Datensatz gefunden werden. Die Übrigen wurden entweder nur in der ersten Befragung genannt ( $N = 128$ , 4,7% aller genannten Kinder) und waren in der zweiten Befragung verschwunden. Oder sie fehlten in der ersten Befragung und waren in der zweiten Befragung neu aufgetaucht ( $N = 63$ ). Wie erwartet ist das Irrtumsrisiko für Kinder also sehr gering, tritt aber gerichtet auf: Kinder werden weitaus häufiger in der ersten, nicht aber in der zweiten Befragung angegeben (vergessen) als umgekehrt nur in der zweiten Befragung berichtet (hinzugefügt). Der Verdacht liegt nahe, dass es sich bei den Hinzugefügten zumindest teilweise um Kinder handelt, die tatsächlich erst nach dem ersten Interviewzeitpunkt in den Haushalt der Befragten eingetreten sind; mehr als die Hälfte von ihnen sind nichtleibliche Kinder der Befragten. Da in der zweiten Befragung das Datum des Einzugs für nichtleibliche Kinder nicht erhoben wurde, lässt sich dies nicht überprüfen. Die Sterbedaten der vier gestorbenen Kinder liegen allerdings sämtlich vor dem ersten Interviewzeitpunkt; diese Kinder waren also vermutlich auch schon zum Zeitpunkt der ersten Befragung gestorben und hätten dort genannt werden müssen.

Abbildung 4.5: Verteilung von Kindschaftsverhältnis und Vitalstatus in der ersten und zweiten Befragung



Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

Tabelle 4.1: Irrtumsrisiko, Fehlergröße und Richtung des Fehlers bei der Rekonstruktion von Kindern, ihren Geburts- und Auszugsdaten

	Leibliche lebendige		Kinder		Gestorbene		Geburtsdaten (Monat und Jahr)		Auszugsdaten (Jahr)		N Kinder			
	<i>n</i>	Irrtums- risiko	<i>n</i>	Irrtums- risiko	<i>n</i>	Irrtums- risiko	<i>n</i>	Irrtums- risiko	<i>n</i>	Irrtums- risiko	<i>n</i>	Irrtums- risiko		
Alle	2.520	3,1	118	28,2	69	32,0	2.308	8,8	986	68,7	678	3,41	1.312	10,8
Männer	1.199	4,6	93	29,0	30	53,3	1.201	13,4	500	71,8	359	3,57	627	15,0
Frauen	1.321	1,2	25	28,0	39	41,0	1.333	4,6	486	65,3	319	3,21	685	7,0
Kohorten: 1930er	677	2,9	31	38,7	37	48,6	701	9,4	545	68,1	371	3,14	312	16,3
1940er	629	3,1	25	40,0	20	35,0	648	11,4	392	70,9	278	3,07	343	12,2
1950er	567	4,1	40	20,0	10	50,0	599	7,1	47	59,6	28		295	8,7
1960er	570	2,1	22	18,2	0		586	7,1	2				131	6,3
1 Kind	369	2,2	17	11,8	5	60,0	374	7,9	125	67,2	84	4,01	391	4,9
2 Kinder	1.165	1,7	37	29,7	15	40,0	1.170	7,5	374	66,8	250	3,00	607	7,2
3 Kinder	499	3,4	29	17,2	24	66,7	508	12,5	237	68,4	237	3,53	184	20,7
4 und mehr Kinder	487	6,6	35	45,7	25	28,0	480	9,0	250	72,8	182	3,56	119	25,5

Irrtumsrisiko: Prozentualer Anteil der abweichenden an allen Angabenpaaren.

Fehlergröße: Mittlere numerische Abweichung ohne Vorzeichen.

Richtung: Prozentualer Anteil der Abweichungen, in denen in der zweiten Befragung ein früheres Datum angegeben wird.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen.

Wie aus Tabelle 4.1 ersichtlich, haben die nichtleblichen und bereits gestorbenen Kinder aus der ersten Befragung ein deutlich höheres Risiko, in der zweiten Befragung nicht erneut angegeben zu werden: Während für leibliche, lebendige Kinder das Risiko lediglich 3,1 Prozent beträgt, liegt es für die 118 nichtleblichen Kinder aus der ersten Befragung bei 28 Prozent. Für bereits gestorbene Kinder fällt dieser Effekt noch massiver aus: Für die 68 in der ersten Befragung bereits gestorbenen Kinder beträgt die Vergessenswahrscheinlichkeit in der zweiten Befragung 47 Prozent, das heißt, fast die Hälfte erscheint in der zweiten Befragung nicht mehr. Für Männer ist das Irrtumsrisiko bei „Standardkindern“ deutlich höher als für Frauen, nicht aber für nichtlebliche Kinder. Kohorteneffekte sind uneindeutig – die Befragten der mittleren Kohorten haben für „Standardkinder“ deutlich höhere Irrtumsrisiken als die älteste und die jüngste Kohorte; nichtlebliche Kinder hingegen werden von der jüngsten und zweitjüngsten Kohorte konsistenter angegeben als von den älteren Kohorten. Für leibliche wie nichtlebliche Kinder steigert eine höhere Kinderzahl das Risiko, eines der Kinder ganz zu vergessen. Kinderseitige und personseitige Faktoren beeinflussen also interaktiv das Irrtums-

Tabelle 4.2: Einfluss personen- und kinderseitiger Faktoren auf die Wahrscheinlichkeit, ein Kind nicht konsistent anzugeben

		Exp(b)
Befragtenseitige Variablen	<i>Kohorte</i> (Referenzkategorie: um 1960)	
	um 1950	1,15
	um 1940	1,16
	um 1930	0,96
	<i>Anzahl Kinder</i> (Referenzkategorie: 1 Kind)	
	2 Kinder	0,96
	3 Kinder	1,61
	4 und mehr Kinder	2,68***
	<i>Geschlecht</i> (Referenzkategorie: weiblich)	
	männlich	2,09***
Kinderseitige Variablen	<i>Kindschaftsverhältnis</i> (Referenzkategorie: Kind leiblich)	
	Kind nichtleiblich	8,85***
	<i>Vitalstatus</i> (Referenzkategorie: Kind lebt noch)	
	Kind gestorben	23,61***
	Konstante	-4,42***
	Gesamtmodell	***
	$\chi^2$	221,94
Freiheitsgrade	9	
7n	2.702	
-2 Log-Likelihood	879,6	

Signifikanzniveau: \*\*\*  $p < 0,01$ .

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen.

risiko, sind aber auch miteinander konfundiert: Die Befragten der älteren Kohorte haben mehr Kinder als die der jüngeren Kohorten (die um 1930 Geborenen geben in der ersten Befragung im Schnitt 2,48 Kinder an, die um 1960 Geborenen nur 1,84 Kinder). Von den nichtleiblichen Kindern werden deutlich mehr von Männern berichtet, und auch Befragte mit mehr Kindern haben mehr nichtleibliche unter diesen: Bei den Personen mit einem oder zwei Kindern in der ersten Befragung ist nur in 5 Prozent der Fälle ein nichtleibliches oder gestorbenes darunter, während es unter denjenigen Personen, die drei oder mehr Kinder haben, schon 26 Prozent sind.

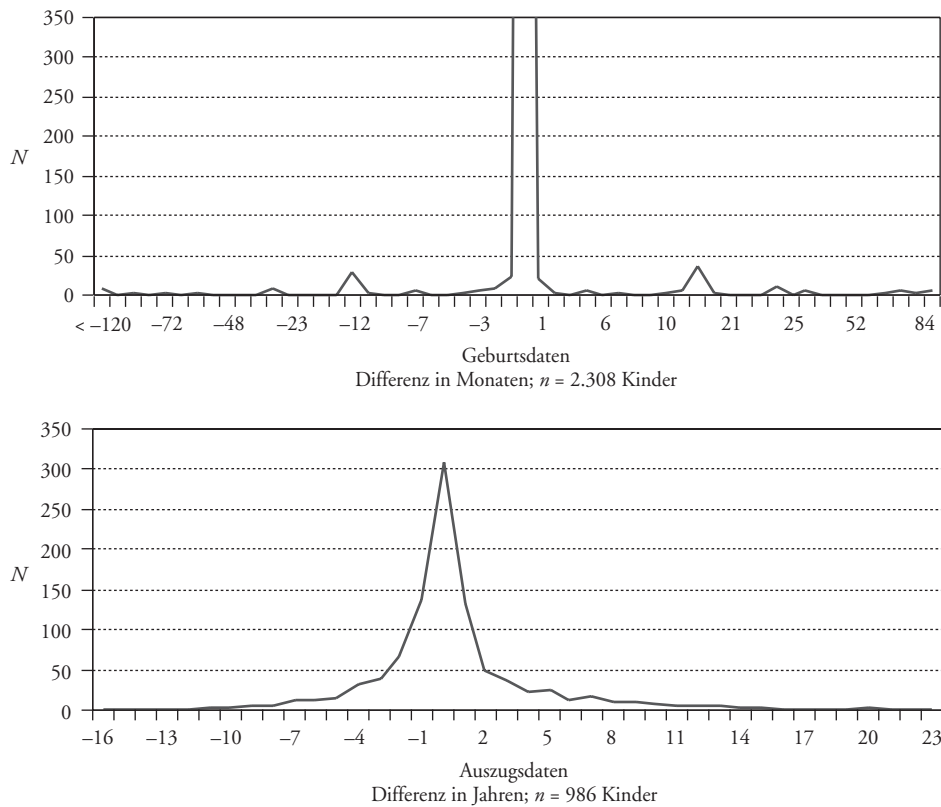
Um abzuschätzen, welche Faktoren nun das Irrtumsrisiko ausschlaggebend steigern oder senken, setze ich verschiedene Befragten- und Informationseigenschaften als unabhängige Variablen in einem multivariaten logistischen Regressionsmodell zur Vorhersage der dichotomen abhängigen Variablen „vergessen/konsistent angegeben“ ein. Die Modellkoeffizienten  $\exp(b)$  können als Erhöhung des Irrtumsrisikos interpretiert werden, das eine Gruppe einer Referenzgruppe gegenüber hat. Als Referenzgruppe setze ich deshalb jeweils die ein, die von der Theorie her bessere Erinnerungsleistung zeigen müssten.

Die Ergebnisse sind in Tabelle 4.2 dargestellt. Das Irrtumsrisiko von Männern ist gegenüber dem von Frauen um mehr als das Doppelte erhöht; die älteren Kohorten hingegen haben kein signifikant erhöhtes Irrtumsrisiko, und die Anzahl der Kinder insgesamt steigert das Irrtumsrisiko erst ab vier und mehr Kindern signifikant um fast das Doppelte. Die kinderseitigen Variablen erhöhen das Irrtumsrisiko beträchtlich: Gegenüber einem noch lebendigen Kind hat ein bereits gestorbenes ein um den Faktor 23 und ein nichtleibliches ein um den Faktor 8 erhöhtes Risiko, etwa 4 Jahre später unterschlagen zu werden. Sowohl Faktoren des Erinnernden als auch der Kinder haben also einen Einfluss darauf, ob Kinder in der zweiten Befragung konsistent wieder angegeben werden.

#### 4.3.2 Datierung von Geburtstagen und Auszügen

In den Vergleich der Geburts- und Auszugsdaten kann ich 2.559 Kinder einbeziehen, die sich einander in den beiden Befragungen zuordnen lassen. Davon liegen in beiden Befragungen für 2.308 Kinder gültige Angaben zu Geburtsmonat und -jahr vor und für 986 gültige Auszugsdaten (die meisten Kinder der jüngsten Kohorten lebten in beiden Befragungen übereinstimmend noch bei ihren Eltern). In Tabelle 4.1 sehen wir das Irrtumsrisiko (als Anteil der abweichenden an allen Angabenpaaren), die Größe des Fehlers (als die mittlere Differenz in Jahren zwischen den abweichenden Daten) und die Richtung des Fehlers (als Anteil aller der abweichenden Angaben an allen, die ein zu frühes Datum angeben). In Abbildung 4.6 ist die Verteilung der numerischen Abweichungen aller Datierungen dargestellt. Daraus geht hervor, dass Ausmaß und Verteilung der Abweichungen die kognitiven Datenkonstruktionsschritte widerspiegeln. Zum Ersten ist das Irrtumsrisiko für das Auszugsdatum deutlich höher als die für das Geburtsdatum: Das exakte Geburtsdatum (Jahr und Monat) wird bei insgesamt 8,8 Prozent der Kinder in der zweiten Befragung verfehlt, für das Auszugsjahr liegt das Irrtumsrisiko hingegen bei 68 Prozent. Zum Zweiten sind die Abweichungen für die Auszugsdaten auch deutlich größer; bei den Geburtsdaten betragen sie im Mittel zwei Monate, und zwei Drittel liegen um weniger als ein Jahr daneben. Die Auszugsdaten hingegen weichen im Schnitt um 2,3 Jahre voneinander ab, und 39 Prozent liegen um weniger als ein Jahr, 75 Prozent weniger als vier Jahre

Abbildung 4.6: Datierung des Geburtstags und des Auszugs – Verteilung der Abweichungen



Positive Werte: Datum in der zweiten Befragung später; negative Werte: Datum in der zweiten Befragung früher.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

auseinander. Beim Geburtsdatum sieht man einen deutlichen Skaleneffekt, der die Unabhängigkeit der Rekonstruktion von Jahren und Monaten belegt. Was die *Richtung* der Abweichungen betrifft, so werden sowohl Geburtsdaten als auch Auszugsdaten nahezu symmetrisch fehldatiert.

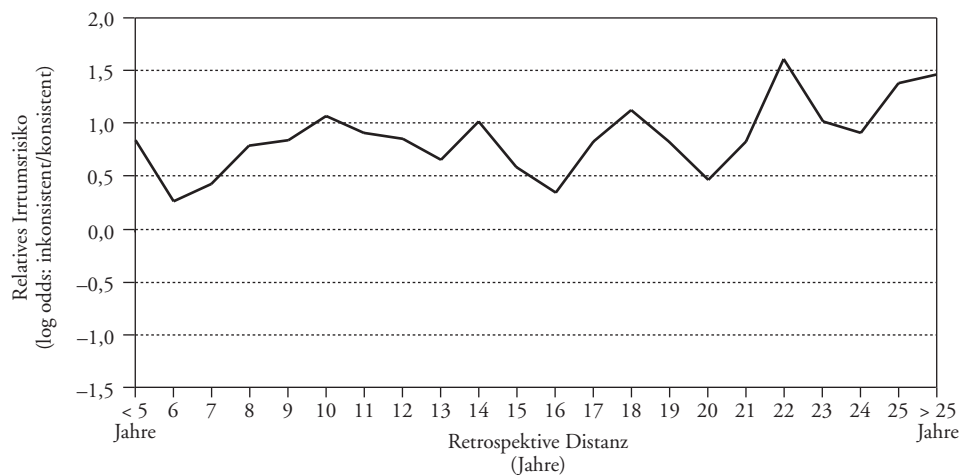
Aus Tabelle 4.1 geht hervor, dass Männer beide Ereignisse mit einem höheren Irrtumsrisiko und etwas größeren Abweichungen datieren. Die beiden ältesten Kohorten irren sich bei Geburtsdaten häufiger als die beiden jüngeren. Für Auszugsdaten können nur die beiden ältesten Kohorten verglichen werden, da zum Zeitpunkt der Befragungen noch die meisten Kinder der beiden jüngeren Kohorten bei ihren Eltern lebten; von diesen beiden hat die Kohorte der um 1930 Geborenen ein etwas geringeres Irrtumsrisiko und einen etwas größeren Fehler. Eine höhere Kinderzahl führt ebenfalls zu einem erhöhten Irrtumsrisiko bei Geburts- und Auszugsdaten, allerdings nicht zu größeren Abweichungen oder einer veränderten Richtung der Abweichungen.



### 4.3.3 Die Datierung von Auszügen und die retrospektive Distanz

Die retrospektive Distanz zu einem Auszug berechne ich als Differenz zwischen dem Datum der zweiten Befragung (in den Jahren 1996 oder 1997) und der Datumsangabe aus der ersten Befragung in Jahren. Der „jüngste“ Auszug liegt vier Jahre zurück<sup>4</sup>, der „älteste“ 40 Jahre. Abbildung 4.7 zeigt das *Irrtumsrisiko* für jüngere und ältere Auszüge, angegeben als *log odds*. Dabei handelt es sich um ein Maß für die Wahrscheinlichkeit, dass von zwei Ereignissen das eine eintritt, im Vergleich zu der Wahrscheinlichkeit, dass das andere Ereignis eintritt. In diesem Fall gibt es das Risiko für die *Inkonsistenz* eines Angabenpaares an und wird errechnet, indem die Wahrscheinlichkeit für das Auftreten des inkonsistenten Falls durch die Wahrscheinlichkeit für das Auftreten des konsistenten Falls geteilt und der natürliche Logarithmus dieses Quotienten berechnet wird. Positive Werte zeigen an, dass inkonsistente Paare häufiger sind als konsistente, und negative Werte zeigen an, dass konsistente Angaben die inkonsistenten überwiegen<sup>5</sup>. Beispielsweise wurden im Zeitraum von sieben bis acht Jahren vor der

Abbildung 4.7: Einfluss der retrospektiven Distanz auf das Irrtumsrisiko für Auszugsdaten



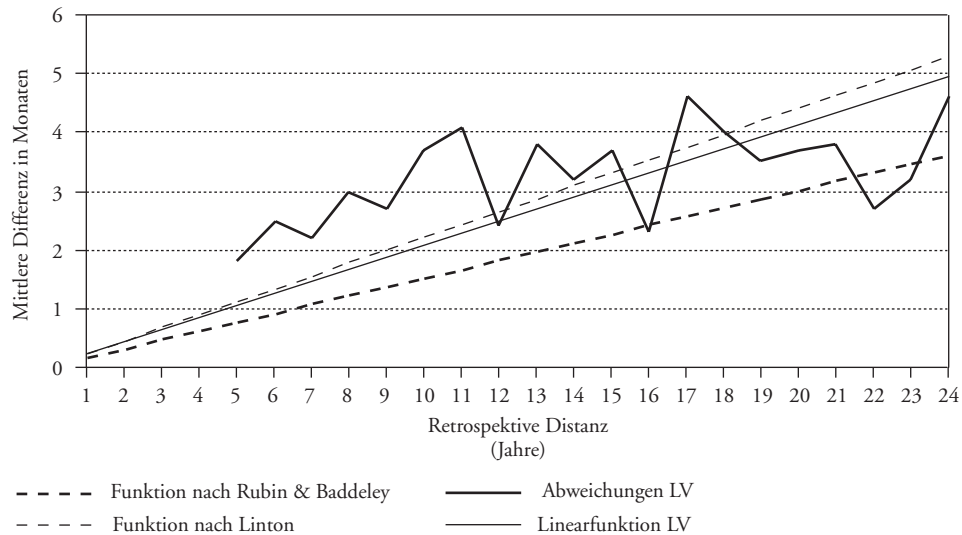
Relatives Risiko (*log odds*) eines Auszugs, abweichend datiert zu werden für jedes Jahr der retrospektiven Distanz.  $n = 986$  abweichend datierte Auszüge. Positive Werte: mehr konsistente Datierungen; negative Werte: mehr inkonsistente Datierungen.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

<sup>4</sup> Dies bedeutet nicht, dass in den drei Jahren vor der zweiten Befragung kein einziges Kind das elterliche Haus verlassen hat. Vielmehr liegt es daran, dass ich nur diejenigen Auszüge in den Vergleich einbeziehen kann, für die in beiden Befragungen Angaben zum Auszugsdatum vorliegen. Kinder, die zwischen den Befragungen das Haus verließen, werden in der ersten Befragung noch als „noch nicht ausgezogen“ verbucht und fallen somit aus der Analyse heraus.

<sup>5</sup> Die Alternative wäre, Wahrscheinlichkeiten anzugeben, die sich als Anteil inkonsistenter Angabenpaare an allen Angabenpaaren berechnen; im angeführten Beispiel wäre dies eine Wahrscheinlichkeit von  $35/51 = 68,62$ . Der Vorteil der logarithmierten *odds* ist, dass darin Information über die Wahrscheinlichkeit eines Ereignisses *relativ zu einem anderen* enthalten ist, was im vorliegenden Fall besonders angemessen erscheint.

Abbildung 4.8: Einfluss der retrospektiven Distanz auf die Fehlergröße für Auszugsdaten



Mittlere numerische Differenz in Monaten nach Jahr des Auszugs und lineare Fehldatierungsfunktionen.  $n = 678$  abweichend datierte Auszüge.

Quelle: Linton, 1975; Rubin & Baddeley, 1989; Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

Befragung von 51 Auszügen 16 konsistent und 35 inkonsistent datiert. Der *odds* für inkonsistente Angaben beträgt also  $35/16 = 2,18$ , was ein *log odds* von 0,78 ergibt.

In Abbildung 4.7 ist der *log odds* für jedes Jahr des retrospektiven Zeitraums positiv, das heißt, zu jedem Zeitpunkt überwiegen inkonsistente Angaben die konsistenten. Bis etwa 20 Jahre nimmt das Irrtumsrisiko bestenfalls schwach zu, erst danach gibt es Anzeichen für eine merkliche Zunahme. Ein Effekt des retrospektiven Intervalls auf die *Verbreitung* von Fehlern ist also vorhanden, nimmt aber diskontinuierliche Formen an.

In Abbildung 4.8 ist nur für die inkonsistenten Datierungen die *Fehlergröße* als die numerische Differenz zwischen den Angaben in Monaten aufgetragen. Im Unterschied zum Irrtumsrisiko nimmt das Ausmaß der Abweichung mit dem retrospektiven Intervall zu: Während Auszüge, die bis fünf Jahre zurückliegen, im Schnitt um 18,9 Monate abweichend datiert werden, sind es für Auszüge, die über 25 Jahre zurückliegen, 59 Monate. Diese Zunahme lässt sich durch die Linearfunktion  $y = 0,2051 x$  beschreiben<sup>6</sup>. Diese Funktion würde also pro verstrichenem Tag 0,21 Tage Zunahme vorhersagen und entspricht damit ungefähr denen, die

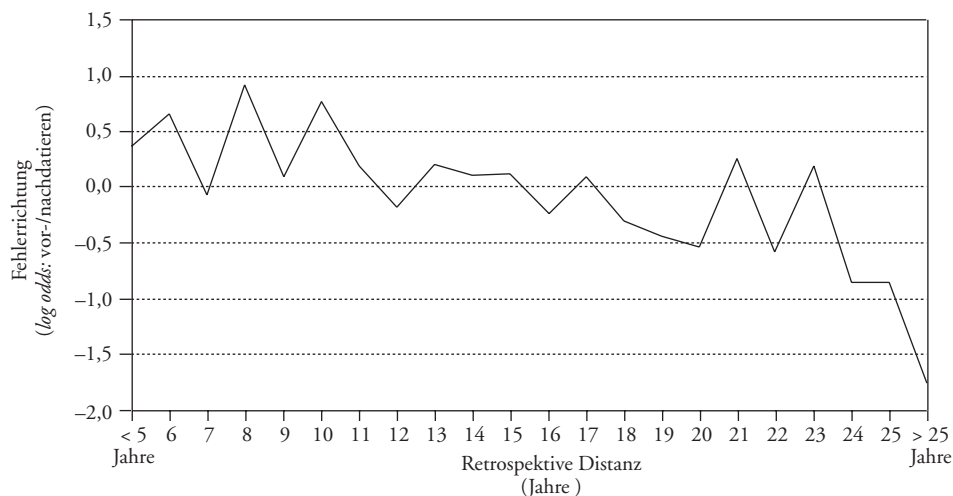
<sup>6</sup> Für retrospektive Distanzen unter fünf Jahren liegen keine Beobachtungen vor. In die Gleichung nehme ich keine Konstante auf, da ich annehme, dass bei 0 Monaten Distanz zum Ereignis (also bei Angaben zur Gegenwart) der Gedächtnisfehler 0 wäre und also auch keine gedächtnisbedingten Abweichungen auftreten würden.

von Linton (1975), Rubin und Baddeley (1989) und Thompson (1982) für Zeiträume bis zu sechs Jahren ermittelt wurden.

Auch die *Richtung* der abweichenden Datierungen hängt mit der Länge der verstrichenen Zeit zusammen, wie Abbildung 4.9 zeigt. Hier sind für alle inkonsistenten Datierungen der *log odds* für Vordatierungen relativ zu Nachdatierungen errechnet und gegen die retrospektive Distanz aufgetragen. Vordatieren bedeutet dabei, dass in der zweiten Befragung ein früheres Datum genannt wird (z.B. 1983 statt 1985). Nachdatieren bedeutet, dass in der zweiten Befragung ein späteres, näher am Befragungszeitpunkt gelegenes Datum genannt wird (z.B. 1988 statt 1985). Ein positiver Wert bedeutet, dass die Vordatierungen häufiger sind, ein negativer Wert bedeutet, dass Nachdatierungen überwiegen. Der Verlauf der Linie zeigt, dass bei „jüngeren“ Auszügen mit kürzerer retrospektiver Distanz die Vordatierungen überwiegen, bei älteren hingegen die Nachdatierungen. Auszugsdaten werden also tendenziell von den „Rändern“ der Referenzperiode hin zu deren Mitte verlagert.

Außerdem hängt die Richtung der Fehldatierung auch damit zusammen, wie alt die Kinder beim Auszug aus dem Elternhaus waren, wie in Abbildung 4.10 deutlich wird. In beiden Kohorten werden die weiter zurückliegenden Auszüge, bei denen die Kinder sehr jung waren, eher nach später verlegt, während die Auszüge, bei denen die Kinder bereits ziemlich alt waren, eher nach früher verlegt wurden<sup>7</sup>. Um das Auszugsalter 21 und 22 Jahre fallen die

Abbildung 4.9: Einfluss der retrospektiven Distanz auf die Fehlerrichtung für Auszugsdaten



Relatives Risiko (*log odds*), in der zweiten Befragung den Auszug vorzudatieren (d.h. ein früheres, weiter vom Befragungszeitpunkt weg gelegenes Datum zu nennen) für jedes Jahr der retrospektiven Distanz.  $n = 678$  abweichend datierte Auszüge. Positive Werte: mehr Vordatierungen (frühere Daten in der zweiten Befragung); negative Werte: mehr Nachdatierungen (spätere Daten in der zweiten Befragung).

<sup>7</sup> Die Auszüge in sehr jungem Alter werden häufig auch Kinder betreffen, die bei einer Trennung des Elternpaares den Haushalt verließen, und stellen in dieser Hinsicht einen Sonderfall dar.

Fehler symmetrisch aus; dieser Wendepunkt markiert also ungefähr den in der Normalbiografie schematisch vorgeschriebenen „Erwartungswert“ für den Auszug eines Kindes. Fehldatierungen rücken in der Regel extremere Auszugsalter an diesen schematischen Erwartungswert heran.

Auszüge, bei denen die Kinder schon älter waren, haben eine kürzere, Auszüge sehr junger Kinder hingegen eine längere retrospektive Distanz. Aus diesem Grund sind Länge der retrospektiven Distanz und Alter beim Auszug miteinander konfundiert. Die Kurven der beiden Kohorten in Abbildung 4.10 verlaufen nahezu identisch, obgleich für die ältere Kohorte die Auszüge im Schnitt etwa zehn Jahre länger zurückliegen. Dies deutet darauf hin, dass die retrospektive Distanz nicht oder weniger ausschlaggebend auf die Fehlerrichtung einwirkt. Um herauszufinden, ob die Veränderungen in der Richtung der Fehldatierungen von der retrospektiven Distanz, dem Alter beim Auszug oder beidem gemeinsam beeinflusst werden, verwende ich die beiden Faktoren in einem multivariaten logistischen Regressionsmodell zur Vorhersage der Wahrscheinlichkeit eines Auszugs, vordatiert zu werden. Die Ergebnisse sind in Tabelle 4.3 dargestellt. Die Modellkoeffizienten  $\exp(b)$  können prinzipiell als Erhöhung des Risikos für das Nach-Früher-Datieren interpretiert werden, das mit der Zunahme einer Einheit der unabhängigen Variable einhergeht. Da allerdings Auszugsalter unter 18 Jahren und über 27 Jahren zusammengefasst wurden und der Nullpunkt der Skala keine Bedeutung hat, ist diese Interpretation der Höhe der Koeffizienten nicht ohne Weiteres möglich. Abgesehen davon nimmt sowohl für alle Befragten zusammen als auch innerhalb der Kohorten die Nei-

Abbildung 4.10: Einfluss des Alters eines Kindes beim Auszug auf die Fehlerrichtung



Relatives Risiko (*log odds*), in der zweiten Befragung ein früheres, weiter vom Befragungszeitpunkt weg gelegenes Datum zu nennen, für unterschiedliche Auszugsalter der Kinder (Angabe aus der ersten Befragung).  $n = 678$  abweichend datierte Auszüge. Positive Werte: mehr Vordatierungen (frühere Daten in der zweiten Befragung); negative Werte: mehr Nachdatierungen (spätere Daten in der zweiten Befragung).

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

Tabelle 4.3: Einfluss der retrospektiven Distanz und des Alters eines Kindes beim Auszug auf das Risiko eines Auszugs, vordatiert zu werden

		Exp(b)
Alle	Retrospektive Distanz	-0,015 n.s.
	Alter des Kindes bei Auszug	0,337***
	Konstante	-1,37***
	Gesamtmodell	***
	$\chi^2$	153,68
	Freiheitsgrade	2
	-2 Log-Likelihood	785,5
	n	678
1930er Kohorte	Retrospektive Distanz	-0,340 n.s.
	Alter des Kindes bei Auszug	0,296***
	Konstante	-0,738 n.s.
	Gesamtmodell	***
	$\chi^2$	90,18
	Freiheitsgrade	2
	-2 Log-Likelihood	423,53
	n	371
1940er Kohorte	Retrospektive Distanz	-0,365 n.s.
	Alter des Kindes bei Auszug	0,328***
	Konstante	-1,13*
	Gesamtmodell	***
	$\chi^2$	54,93
	Freiheitsgrade	2
	-2 Log-Likelihood	329,75
	n	278

Signifikanzniveau: \*  $p < 0,1$ ; \*\*\*  $p < 0,01$ ; n.s. = nicht signifikant.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

gung zum Vordatieren mit jedem Jahr des Kindesalters hochsignifikant zu, die retrospektive Distanz hingegen übt keinen eigenen Einfluss (mehr) aus. Die Fehlerrichtung wird also vor allem durch den „Druck“ eines schematisch vorgeschriebenen Auszugsalters für Kinder hervorgerufen.

#### 4.4 Diskussion

##### *Zusammenfassung: Gedächtnisfehler und ihre Folgen*

Die Erinnerung an Kinder hat sich wie erwartet für leibliche Kinder als sehr stabil über die Zeit erwiesen, während nichtleibliche und bereits gestorbene Kinder in der längeren Rückschau ein beträchtliches Risiko haben, unterschlagen zu werden. Die Angaben von Männern sind in fast jeder Hinsicht weniger konsistent als die von Frauen, während die Angaben älterer Kohorten sich nicht eindeutig von denen jüngerer Kohorten unterscheiden. Eine erhöhte Kinderanzahl erhöht das Risiko, eines gänzlich zu unterschlagen, ab vier Kindern. Der Datierungsfehler hingegen wird von der Kinderzahl nicht beeinflusst.

Geburtsdaten werden deutlich konsistenter erinnert als Auszugsdaten. Bei Geburtsdaten erfolgt die Rekonstruktion meistens durch direkten Zugriff auf ein Zeitetikett; das Ereignis Kindsgeburt ist unmissverständlich und kaum interpretationsanfällig; es wird intensiv geübt; es ist in den temporalen Strukturen des autobiografischen Gedächtnisses vielfältig verknüpft und reich an temporalen Abruf-Hinweisreizen. Bei Auszügen hingegen muss so gut wie immer erst das fragliche Ereignis und dann sein Datum rekonstruiert werden. Der Auszug eines Kindes aus dem Elternhaus ist ein weniger klar und eindeutig umschriebenes Ereignis; Auszüge werden weniger intensiv geübt; sie sind in geringerem Maße in die zeitlichen Strukturen des autobiografischen Gedächtnisses eingebunden und verfügen über weniger gute temporale Abruf-Hinweisreize. Alle Datierungsfehler fallen symmetrisch aus; Auszüge werden aber in Abhängigkeit vom tatsächlichen Sachverhalt fehldatiert und in der längeren Rückschau eher in die Nähe eines „Standardalters“ von etwa 21 verschoben.

Die in der psychologischen Gedächtnisforschung nachgewiesene Neigung, den eigenen Lebenslauf in der Rückschau durch selektives Vergessen und Umdeuten als kontinuierlicher, kohärenter und konventioneller zu rekonstruieren, als es den tatsächlichen Gegebenheiten entspricht, macht sich auch bei der Erhebung im Bereich Nachwuchs und Familie bemerkbar. Die Folge für die zu einem späteren Zeitpunkt erhobenen Daten ist, dass sich die Lebensverläufe der Befragten als konventioneller und stärker im Einklang mit normativen Vorgaben einer Standardbiografie darstellen: „Nichtstandardkinder“ werden aus den Lebensläufen getilgt, ebenso die jung verstorbenen, und die Kinder verlassen das Haus eher zu einem normativ vorschriftsmäßigen Zeitpunkt. Fehldatierungen rücken in der Regel extremere Auszugsalter an diesen schematischen Erwartungswert heran. Dieser liegt bei etwa 21 oder 22 Jahren und reflektiert damit tatsächlich in etwa das mittlere Auszugsalter für ostdeutsche Kinder des fraglichen Zeitraums (Hullen, 1995). Damit soll nicht gesagt sein, dass normative Erwartungswerte immer die realen Gegebenheiten reflektieren. Zumal das tatsächliche Auszugsalter über diesen Zeitraum hinweg gestiegen ist, wäre es interessant herauszuarbeiten, ob sich dadurch auch Erwartungswerte verändern.

#### *Je länger die Rückschau, desto schlechter die Angaben?*

Die Analysen haben gezeigt, dass zwar eine spezifische Gedächtnisfehlerkomponente in den retrospektiven Daten besteht, dass die retrospektive Distanz aber verschiedene Aspekte des Fehlers in verschiedenem Ausmaß und auch auf verschiedenen Wegen beeinflusst. Das *Irrtumsrisiko* nimmt mit der Länge der retrospektiven Distanz nur geringfügig zu; wer das Auszugsjahr seiner Kinder stabil weiß oder sehr gut rekonstruieren kann, kann dies fünf Jahre danach ebenso gut wie 20 Jahre danach. Dies könnte den ersten Schritt des Konstruktionsprozesses widerspiegeln, bei dem das zu datierende Ereignis rekonstruiert werden muss: Diejenigen, die in beiden Befragungen jeweils dasselbe Ereignis im Sinn haben, unterliegen nur noch dem vergleichsweise geringen Fehler bei der zeitlichen Einordnung dieses Ereignisses. Diejenigen, die hingegen aus zwei oder mehreren Ereignissen, die als Auszug gelten können, eines „auswählen“ müssen, produzieren die größeren Abweichungen. Diejenigen allerdings, die abweichende Angaben machen, machen tendenziell für länger zurückliegende Auszüge *größere* Fehler. Die Linearfunktion, die diese Zunahme der Größe der Abweichungen beschreibt, ähnelt derjenigen, die auch für wesentlich kürzere retrospektive Distanzen und Perioden gefunden wurden.

Da die Vergessens- bzw. Fehldatierungskurve auf Daten einer Wiederholungsbefragung beruht, müssen die Kurven anders interpretiert werden als beim Vergleich von Befragtenangaben mit externen Aufzeichnungen (vgl. Abb. 3.5): Eine lineare Zunahme der Größe der Abweichungen steht nicht für eine lineare Zunahme der Erinnerungsfehler (Abb. 3.5a) – eine *lineare* Zunahme des Erinnerungsfehlers würde sich in *gleichbleibenden* Abweichungen der wiederholten Angaben niederschlagen. Hinter einer linearen Zunahme könnte sich die exponentielle Funktion aus Abbildung 3.5c verbergen. Das hieße, dass die Kombination aus Abweichungen bei der Ereignisrekonstruktion und Datierungsfehlern in den Fällen, in denen abweichende Angaben gemacht werden, zu einer *exponentiellen* Zunahme der Abweichungen mit zunehmender retrospektiver Distanz führt. Ebenso steht das gleichbleibende Irrtumrisiko möglicherweise für eine lineare Zunahme des Erinnerungsfehlers über die Zeit.

Mit der retrospektiven Distanz ändert sich außerdem die *Richtung*, in die Abweichungen auftreten: Das Auszugsdatum für die am längsten zurückliegenden Auszüge wird bevorzugt früher erinnert, für die am kürzesten zurückliegenden Auszüge hingegen bevorzugt später. Dies aber stellt keinen Effekt der verstreichenden Zeit dar, sondern einen Effekt des Alters der Kinder beim Auszug: Je stärker dieses Alter nach oben oder unten von einem normativen Erwartungswert von 21 oder 22 Jahren abweicht, desto eher werden Auszugsdaten an diesen herangerückt. Es handelt sich also um einen Faktor, der mit dem Ereignis des Auszugs bzw. dem erhobenen Lebenslauf selbst assoziiert ist: der Grad seiner Entsprechung mit den schematischen Erwartungen an Lebensverläufe (der Eltern und der Kinder). Die Länge der retrospektiven Distanz hingegen übt keinen eigenen Einfluss auf die Fehlerrichtung aus.

## Kapitel 5

„Waren Sie im Dezember 1989 erwerbstätig?“

### Rekonstruktion und Reliabilität von Erwerbsverläufen im Transformationsprozess

#### 5.1 Fragestellung und Vorgehensweise

Erwerbsbiografien sind für das Verständnis von Lebensläufen von zentraler Bedeutung. Erwerbsverläufe fungieren unter anderem als Mechanismus der Allokation von Personen zu Positionen und Ressourcen und vollziehen sich im institutionellen Gefüge einer Gesellschaft, sodass Veränderungen auf der Makroebene sich in individuellen Verläufen widerspiegeln. Außerdem können Erwerbsverläufe als Ergebnisse der Entscheidungen und Handlungsstrategien der Personen betrachtet werden, mit denen sie unter den gegebenen Bedingungen diesen zentralen Bereich ihres Lebens gestalten und ihre Ziele verfolgen. Um die Folgen von Veränderungen in Arbeitsmarktbedingungen und Beschäftigungssystemen und den Ertrag unterschiedlicher berufsbiografischer Entscheidungen unter bestimmten Rahmenbedingungen einzuschätzen, ist es deshalb wichtig, dass die Erwerbsverläufe sowohl in ihren übergeordneten zeitlichen Eigenschaften als auch in ihrer Episodenstruktur korrekt abgebildet werden.

In diesem Kapitel betrachte ich die Reliabilität der Angaben zur Erwerbsbiografie aus der Wiederholungsbefragung der *Deutschen Lebensverlaufstudie* in Ostdeutschland (LV-Ost). Die beiden Befragungen, die in einem Abstand von vier bis fünf Jahren durchgeführt wurden, decken einen Zeitraum von ungefähr zwei Jahren zwischen Dezember 1989 bis Mitte 1992 ab. Da es sich hierbei um Ereignisdaten handelt, stellt die angemessene Beschreibung der Abweichungen eine besondere Herausforderung dar.

Die folgenden Fragen werden im Vordergrund stehen:

- Wie verändert sich die Rekonstruktion der Ereignissequenzen zwischen den beiden Befragungen, und wie lassen sich diese Veränderungen angemessen als Folge von verändert ablaufenden Gedächtnisprozessen beschreiben?
- Wie sieht die erinnerungsspezifische Fehlerkomponente für die zeitlichen Informationsdimensionen der Sequenzen wie Inzidenz, Frequenz, Dauer und Timing aus?
- Was sind die Folgen der Abweichungen für die Stabilität deskriptiver und explanatorischer inhaltlicher Analysen?
- Fallen die Fehler in den edierten Daten geringer aus, und welche Information profitiert besonders von der Editionsarbeit?

Ich werde in Abschnitt 5.2.1 zunächst darstellen, wie Erwerbsverläufe in der LV-Ost erhoben werden, und – im Sinne eines partiellen Datengenerierungsmodells – die kognitiven Mechanismen herausarbeiten, die die Befragten bei der Rekonstruktion der Erwerbssequenzen einsetzen. Auf dieser Grundlage werde ich Erwartungen über die Abweichungen zwischen den



Befragungen ableiten und dabei besonders die Erinnerungsbedingungen der Transformationsperiode berücksichtigen. Ich werde mich auf die Erhebung von Phasen der *Erwerbstätigkeit*, *Arbeitslosigkeit* und *Nichterwerbstätigkeit* und die Übergänge zwischen ihnen konzentrieren. Damit werden drei qualitativ unterschiedliche Zustände der Arbeitsmarktbeteiligung erfasst. Eine Person partizipiert entweder aktiv am Arbeitsmarkt und ist in der Hauptsache erwerbstätig (Erwerbstätigkeit), sie würde gern partizipieren, steht dem Arbeitsmarkt auch zur Verfügung und ihre Hauptbeschäftigung ist das Betreiben einer Rückkehr zu bzw. Neuaufnahme von aktiver Partizipation (Arbeitslosigkeit), oder sie hat sich zu Gunsten einer anderen Aktivität wie Familienarbeit, Ruhestand, Krankheit oder Ausbildung aktuell oder auf Dauer vom Arbeitsmarkt zurückgezogen (Nichterwerbstätigkeit).

In Abschnitt 5.3.2 nehme ich eine qualitative Analyse der Abweichungen anhand ausgewählter Einzelfälle vor. In den Abschnitten 5.3.3 und 5.3.4 untersuche ich, wie die systematische Gedächtnisfehlerkomponente für die Abbildung folgender Informationsdimensionen der Erwerbssequenz aussieht:

- *Inzidenz* und *Frequenz* von Episoden und Übergängen und die *Zusammensetzung* der Sequenz,
- Anzahl der Monate, die die Befragten in verschiedenen Zuständen verbrachten (*Dauerverteilungen*),
- *Art und Timing des ersten Übergangs* innerhalb des Zeitfensters.

Ich werde den Fehler differenziert im Hinblick auf *Verbreitung*, *Ausmaß* und *Richtung* beschreiben und in Abschnitt 5.3.5 die Interdependenz der Fehler bei Inzidenz bzw. Frequenz einerseits und der Dauerverteilung andererseits untersuchen. In Abschnitt 5.3.7 werde ich die Auswirkungen der Erinnerungsfehler auf die Ergebnisse deskriptiver und explanatorischer inhaltlicher Analysen illustrieren, und in Abschnitt 5.3.8 werde ich prüfen, ob sich die Fehler durch die Datenedition verringert haben.

## 5.2 Datengenerierung und Hypothesen

### 5.2.1 Die retrospektive Erhebung von Erwerbsgeschichten in der LV-Ost

*Die Studie „Lebensverläufe und historischer Wandel in der ehemaligen DDR“ (LV-Ost)*

In der ersten Befragung wurden zwischen September 1991 und August 1992 insgesamt 2.331 Interviews mit bevölkerungsrepräsentativen Stichproben der Geburtsjahrgänge 1929–31, 1939–41, 1951–53 und 1959–61 durchgeführt. Die Teilnehmer, die einer Rekontaktierung zugestimmt hatten ( $N = 1.394$ ), wurden in den Jahren 1996 und 1997 erneut befragt, um die Veränderungen seit der ersten Befragung nachzuzeichnen. Die Erwerbsgeschichte wurde dabei nicht erst ab dem Zeitpunkt des ersten Interviews erhoben, sondern schon ab Dezember 1989. Zwar sind Kalenderdaten sehr schlechte Abruf-Hinweisreize (Barsalou, 1988; Brewer, 1988; Wagenaar, 1986). Aber da im Monat davor, am 9. November 1989, der Fall der Berliner Mauer stattfand, ist dieses Datum mit einem historischen Grenzstein fest verbunden, der einen grundlegenden Wandel in zahlreichen Lebensumständen und Rollen der Befragten markiert.

Durch diese erhebungstechnische Entscheidung liegen zu den Jahren vom Mauerfall bis zum Zeitpunkt der ersten Befragung von 1.049 Befragten der Geburtskohorten 1939–41,

Abbildung 5.1: Das in der LV-Ost zweifach abgedeckte Zeitfenster und gedächtnisrelevante Aspekte der verstreichenden Zeit

	Zweite Befragung					Erste Befragung																	
	Zusätzlich verstrichene Zeit					Zeitfenster																	
Jahr	97	96	95	94	93	<b>92</b>	<b>91</b>	<b>90</b>	89	88	87	86	85	84	83	82	81	80	79	78	77	76	...
Retrospektive Distanz (Jahre)																							
Erste Befragung						<b>0</b>	<b>1</b>	<b>2</b>	3	4	5	6	7	8	9	10	11	12	13	14	15	16	...
Zweite Befragung	0	1	2	3	4	<b>5</b>	<b>6</b>	<b>7</b>	8														
Ungefähres Alter (Jahre)																							
Kohorte 1939–41	57	56	55	54	53	<b>52</b>	<b>51</b>	<b>50</b>	49	48	47	46	45	44	43	42	41	40	39	38	37	36	...
Kohorte 1949–51	47	46	45	44	43	<b>42</b>	<b>41</b>	<b>40</b>	39	38	37	36	35	34	33	32	31	30	29	28	27	26	...
Kohorte 1959–61	37	36	35	34	33	<b>32</b>	<b>31</b>	<b>30</b>	29	28	27	26	25	24	23	22	21	20	19	18	17	16	...

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

1949–51 und 1959–61<sup>1</sup> zwei Erwerbssequenzen vor (siehe Abb. 5.1). In der ersten Befragung wurden alle Teilnehmer persönlich befragt, in der zweiten wurden etwa 80 Prozent der Befragungen telefonisch durchgeführt<sup>2</sup>.

In der LV wird die Erwerbsgeschichte als zeitlich geordnete Abfolge von Erwerbsepisoden oder „Jobs“ erfasst. Ein neuer „Job“ beginnt immer dann, wenn einer von vier Übergängen eintritt: Wechsel des Betriebs, Veränderung der Arbeitszeit, des Einkommens, der Tätigkeit oder der beruflichen Stellung innerhalb desselben Betriebs. Zeiträume ohne eine Erwerbstätigkeit können in einem von sechs Lückenzuständen verbracht werden: *Arbeitslosigkeit* und *Nichterwerbstätigkeit* (Ausbildung, Hausfrau, Erziehungsurlaub, Rente und Sonstiges)<sup>3</sup>.

#### Daten- und Datenfehlergenerierung

Die kontinuierliche Abfolge von Erwerbsepisoden und Lücken wurde erhoben, indem der Teilnehmer zuerst gefragt wurde, ob er zu Beginn des fraglichen Zeitausschnitts – in der ers-

<sup>1</sup> Die Befragten der 1929–31 geborenen Kohorte waren zum Zeitpunkt der zweiten Befragung, bis auf wenige Ausnahmen im Rahmen flächendeckender Regelungen, die älteren Arbeitnehmern den Übertritt in den vorgezogenen Ruhestand ermöglichten, aus dem Erwerbsleben ausgeschieden. Deshalb wurden diese Personen nicht noch ein zweites Mal zu ihren Erwerbsverläufen befragt.

<sup>2</sup> Zum Zeitpunkt der ersten Befragung war die Telefondichte in den neuen Ländern noch nicht hoch genug, um eine für die Gesamtbevölkerung repräsentative Telefonstichprobe befragen zu können.

<sup>3</sup> Da reale Lebensverläufe zum Teil weniger eindeutig zwischen diesen drei abgrenzbaren Zuständen verlaufen, wird in der LV die Erwerbsgeschichte als ein zweifacher Pfad betrachtet: ein institutionell stark geregelter Hauptpfad von klar abgrenzbaren Erwerbstätigkeiten, die sich mit anderen Vollzeit-Hauptbeschäftigungen abwechseln, und einen weniger konventionellen und institutionell geregelten, in dem unregelmäßige und Nebenerwerbstätigkeiten, Teilzeitausbildungen usw. parallel zum Hauptpfad verlaufen (können). Der Nebenpfad wird gesondert ebenfalls erhoben; diese Angaben werde ich im Folgenden nicht systematisch betrachten.

ten Befragung nach Abschluss der allgemein bildenden Schule bzw. der ersten Ausbildung, in der zweiten Befragung im Dezember 1989 – erwerbstätig war (vgl. Abb. 3.9). War dies der Fall, wurden über diese Erwerbstätigkeit zahlreiche Detailinformationen erhoben (u.a. Einkommen, berufliche Stellung und Rechtsform des Betriebs) sowie Start- und Enddatum der Episode monatsgenau erfasst. Wenn die Erwerbstätigkeit nicht bis zur Gegenwart reichte, wurde gefragt, ob direkt im Anschluss eine weitere Erwerbstätigkeit aufgenommen wurde. War dies der Fall, wurde erfragt, ob sich bei diesem Wechsel der Betrieb, die Arbeitszeiten, die berufliche Stellung oder das Gehalt geändert hatten und danach auch für diese Erwerbsepisode Details und Start- und Enddatum erhoben. War dies nicht der Fall, wurden dem Befragten die sechs „Lückenaktivitäten“ zur Beschreibung angeboten: „Arbeitslos – in Ausbildung/Weiterbildung – Hausfrau – Erziehungsurlaub – Vorruhestand – Sonstiges“. Dann wurde für diesen Zustand das Enddatum festgestellt, und wenn die Episode nicht bis zur Gegenwart reichte, wurde gefragt, ob sich eine Erwerbstätigkeit oder eine weitere Lücke angeschlossen hatte. Dieser Vorgang wiederholte sich so lange, bis der gesamte Zeitraum bis zum Interviewdatum vollständig abgedeckt war.

In Abbildung 5.2 ist dargestellt, welche kognitiven Schritte der Befragte bei diesem Vorgehen vollführen muss. Im ersten Schritt muss er die erwerbs- und arbeitsmarktbezogenen Aktivitäten, Rollen, Orte oder Personen rekonstruieren und sie einer der angebotenen Kategorien *zuordnen*. Dazu gehört auch die Interpretation dieser angebotenen Zustände, das heißt, er muss sich darüber klar werden, was beispielsweise mit „arbeitslos“ gemeint ist. Als zweites muss er *entscheiden*, ob in der Zwischenzeit eine *Veränderung* in diesen Aktivitäten und Rollen, Orten oder Personen eintrat, die einen Wechsel herbeiführte und wenn ja, angeben, was dies für ein Wechsel war. Im dritten Schritt muss er für Anfang und Ende dieser rekonstruierten Episoden bzw. für den rekonstruierten Übergang das kalendarische *Datum* auf den Monat genau *rekonstruieren*.

Bei allen drei Schritten kann der Befragte in der zweiten Befragung zu einem anderen Rekonstruktionsergebnis kommen als in der ersten: Bei Schritt 1 kann er einen anderen Zustand aus dem Zustandsraum der LV auswählen, entweder weil seine Erinnerung an die zentralen Aktivitäten und Rollen sich verändert hat oder weil er die angebotenen Zustände anders interpretiert. Bei Schritt 2 kann er abweichend rekonstruieren, ob und wie die Aktivitäten und Rollen sich veränderten. Bei Schritt 3 kann die Rekonstruktion aus temporalen Abruf-Hinweisreizen oder *order codes* zu abweichenden Monats- oder Jahresangaben für das Ende einer und den Anfang der nächsten Episode führen.

Wenn Abweichungen auf einer früheren Stufe geschehen, sind auch für die folgenden Stufen Abweichungen vorprogrammiert. Wer beispielsweise gleich zu Beginn einen Zustand nicht mehr als Arbeitslosigkeit, sondern als Hausfrau bezeichnet, hat ein erhöhtes Risiko, auch das Ende dieses Zustands anders zu rekonstruieren. Oder wenn ein innerbetrieblicher

Abbildung 5.2: Datengenerierungsprozess bei der Erhebung der Erwerbsgeschichte in der LV-Ost

---

Schritt 1	<i>Episode benennen</i> : Erwerbs- und arbeitsmarktbezogene Aktivitäten, Rollen, Orte und Personen rekonstruieren und einer der Kategorien zuordnen
Schritt 2	<i>Übergang rekonstruieren</i> : Veränderungen in den Erwerbs- und arbeitsmarktbezogenen Aktivitäten, Rollen, Orte und Personen rekonstruieren
Schritt 3	<i>Episode/Übergang datieren</i> : Monat und Jahr des Episodenanfangs und -endes rekonstruieren

---

Wechsel unterschlagen wird und erst ein Wechsel in einen anderen Betrieb, der Monate später stattfand, die Erwerbsepisode beendet, fällt auch die Datierung dieser Übergänge mit höherer Wahrscheinlichkeit abweichend aus.

### 5.2.2 Annahmen I: Was wird abweichend rekonstruiert?

Identität, soziale Rollen und Status, Teilhabe an gesellschaftlichen Angeboten, Ziele, Motive und Lebenspläne werden in industrialisierten Gesellschaften innerhalb und in Bezug auf den Erwerbsbereich entwickelt. Die Erwerbsgeschichte der meisten Menschen dürfte deshalb in hohem Maße organisiert und als ein den subjektiven und normativen Entwicklungslogiken angepasstes Narrativ repräsentiert sein, das die Gegenwart als Folge der Vergangenheit etabliert. Für die Episoden und Übergänge, die als fester Bestandteil dieses Erwerbsnarrativs repräsentiert sind, fällt die Rekonstruktion deshalb sehr stabil aus.

Die in der LV erhobenen Episoden entsprechen dabei *Lebensabschnitten* im Sinne von Conway und Pleydell-Pearce (2000). Lebensabschnitte im Bereich Arbeit und Erwerbsleben werden durch die erwerbs- und arbeitsmarktbezogenen sozialen Rollen, Aktivitäten, Orte und Personen definiert, und Übergänge zwischen Lebensabschnitten entstehen, wenn diese sich in persönlich bedeutsamer und folgenschwerer Weise verändern.

Lebensabschnitte werden am ehesten von Erwerbstätigkeiten als Norm- oder Standardaktivitäten gebildet (Conway, 1996). Phasen der Arbeitslosigkeit hingegen definieren sich vor allem negativ als Abwesenheit von Erwerbstätigkeit und weisen keine eigenständigen Rollen und Aktivitäten auf. Da sie darüber hinaus noch negativ stigmatisiert sind, haben sie in der biografischen Sinnstruktur des Erwerbsverlaufs meistens keinen Platz (Elias, 1997; Mutz, 1996; Paull, 2002). Insofern erwarte ich für sie ein hohes Risiko, instabil erinnert zu werden. Am ehesten bringen Übergänge zwischen Zeiten mit und ohne Erwerbstätigkeit (also Arbeitslosigkeit oder Nichterwerbstätigkeit) sowie Wechsel zwischen Betrieben markante Veränderungen der sozialen Rollen und Aktivitäten mit sich, die den Grenzen von Lebensabschnitten entsprechen. Übergänge zwischen Arbeitslosigkeit und Nichterwerbstätigkeit sowie zwischen Erwerbsepisoden innerhalb desselben Betriebs hingegen sind weniger markant und werden vermutlich eher abweichend rekonstruiert.

### 5.2.3 Annahmen II: Allgemeine und „ostspezifische“ Veränderungen

Durch die autobiografische Rekonstruktion ist eine Tendenz zu erwarten, die Erwerbsverläufe aus der längeren retrospektiven Distanz weniger wechselhaft und konventioneller zu rekonstruieren. Dies bedeutet zunächst eine Verringerung von Übergängen und Episoden und eine Neigung, mehr Standardaktivität (Erwerbstätigkeit) und weniger Arbeitslosigkeit und Nichterwerbstätigkeiten zu berichten.

In das abgedeckte Zeitfenster fällt der im Rahmen der nationalen Vereinigung vollzogene umwälzende Transformationsprozess, bei dem das planwirtschaftliche Wirtschafts- und Beschäftigungssystem der DDR durch das marktwirtschaftliche System der BRD ersetzt wurde. Durch den historisch einmalig rasanten und weit reichenden Zusammenbruch des Arbeitsmarkts, in dessen Verlauf das Arbeitsplatzangebot um 40 Prozent sank, und durch die an-

haltend schwierigen Arbeitsmarktkonstellationen, die von flächendeckender Langzeitarbeitslosigkeit und wirtschaftlicher Stagnation gekennzeichnet sind, stellt sich der Erinnerungskontext in der zweiten Befragung anders dar als in der ersten.

Goedicke (2000) identifiziert vier Phasen der Transformation (siehe Abb. 5.3), von denen die ersten beiden in etwa das Zeitfenster des Vergleichs abdecken. Auf ein halbes Jahr, in dem für die Arbeitnehmer vergleichsweise wenige Veränderungen eintraten, folgte ein Zeitraum von gut einem Jahr bis Ende 1991, in dem Grundsicherung und Institutionentransfer einen massiven Abbau von Personal (z.T. abgefedert durch arbeitspolitische Maßnahmen wie Ruhestand, Arbeitsbeschaffungsmaßnahme oder Kurzarbeit) mit sich brachten. Dieser vollzog sich in einem Kontext von massenhaften Betriebsschließungen, während die weiter bestehenden Betriebe durch Abschaffung und Ausgründung von Betriebsteilen radikal umstrukturiert wurden und meistens Rechtsform und/oder Träger wechselten. Ab etwa 1992 folgten die Phasen der „betriebswirtschaftlichen Konsolidierung“ und „Modernisierung und Normalisierung“, die den Übergang von einer akuten wirtschaftsstrukturellen Umbruchsituation in eine langfristige Beschäftigungs- und Strukturkrise markieren. Ich erwarte, dass diese Veränderungen des Rekonstruktionskontextes sich in der Wahrnehmung von Lebensverläufen zwischen den Befragungen in den Abweichungen der Angaben zu Arbeitslosigkeit widerspiegeln und dazu führen, dass 1996, wenn Arbeitslosigkeit individuelles Schicksal und kollektive „Standarderfahrung“ geworden ist, Zustände bereitwilliger als Arbeitslosigkeit eingeschätzt werden als 1992.

Auch dass Frauen sich an Erwerbstätigkeit und Arbeitslosigkeit schlechter erinnern als Männer (Dex & McCulloch, 1997; Elias, 1997; de Graaf & Wegener, 1989; Paull, 2002; Thélot, 1990) und stärker dazu neigen, Arbeitslosigkeit in Nichterwerbstätigkeiten umzudefinieren (Paull, 2002), ist für die ostdeutschen Befragten nicht im selben Maß zu erwarten. Das DDR-Beschäftigungssystem war auf lebenslange Vollbeschäftigung ausgerichtet, hatte eine der höchsten Partizipationsraten der gesamten Welt und verstand sich als zentrale Achse des gesamten Lebens der Bürger beider Geschlechter und aller Altersgruppen, deren Wohn-,

Abbildung 5.3: Phasen des ostdeutschen Transformationsprozesses

---

*Dezember 1989–Juni 1990 – Wirtschaftspolitische Reorientierung in der DDR-Reformphase*

Keine massenhaften inner- oder zwischenbetrieblichen Mobilitätsprozesse außer Westmigration jüngerer Arbeitskräfte

*Juli 1990–Dezember 1991 – Institutionentransfer und betriebliche Grundsicherung*

Massenarbeitslosigkeit, nur entlastet durch Maßnahmen staatlicher Arbeitsmarktpolitik (vor allem Verringerung der Arbeitsplatznachfrage, Arbeitsbeschaffungsmaßnahmen) und Migration nach Westen  
Radikale rechtliche und strukturelle Neugestaltung der Arbeitsmarkt- und Betriebsstrukturen

*Januar 1992–Juni 1994 – Betriebswirtschaftliche Konsolidierung*

Vergleichsweise verlangsamte und selektive Umgestaltung, zunehmende Etablierung der neu eingeführten Institutionen und Strukturen

*ab Juli 1994 – Sichtbare Modernisierungserfolge*

Verlangsamung der Privatisierungen und des Personalabbaus, Institutionentransfer der Verwaltung und Wirtschaftsverbände abgeschlossen, Auslauf der stark ausgeweiteten beschäftigungssichernden Maßnahmen der Arbeitsmarktpolitik

---

Quelle: Goedicke, 2000.

Familien- und Freizeitverhalten in Bezug auf die Erwerbsarbeit und den Betrieb gestaltet wurden (Diewald, 1999; Goedicke, 2000). Beide Geschlechter nahmen also in ähnlichem Umfang am Erwerbsleben teil und wurden in ähnlicher Weise im Hinblick auf ihre arbeitsmarkt- und erwerbsbezogenen Rollen und Erwartungen sozialisiert. Noch Jahre nach der Wende ist die Erwerbsorientierung ostdeutscher Frauen wesentlich stärker als die westdeutscher Frauen; ein Rückzug in die Hausfrauenrolle fand nicht statt (Berger, 1995; Harre & Schmidt, 1995; Nickel, 1994). Deswegen nehme ich an, dass die ostdeutschen Frauen auch weniger bereit sind, Zeiträume unfreiwilliger Arbeitsmarktaktivität als Nichterwerbstätigkeit (Hausfrau, Erziehungsurlaub usw.) zu bezeichnen und dass die Geschlechterunterschiede nicht oder in geringerem Maß auftreten.

Die Bedingungen des Transformationsprozesses machen auch Unterschiede zwischen den befragten Kohorten weniger wahrscheinlich. Paull (2002) belegte, dass die frühen Phasen der Erwerbstätigkeit, in denen die Befragten ihren Platz im Beschäftigungssystem suchten und sich orientierten, schlechter rekonstruiert werden. Dies spielt für die jüngste Befragtenkohorte, die Anfang der 1990er Jahre um die 30 Jahre alt war, am ehesten noch eine Rolle. Im Transformationsprozess in Ostdeutschland veränderte sich das System jedoch rasch auf einen festgeschriebenen Zielzustand hin und schuf mehr oder weniger zeitgleich für alle in ihm lebenden Generationen ähnlich unsichere und wechselhafte Bedingungen (Diewald, 1999). Insofern haben auch ältere Personen wechselhafte und unregelmäßige Erwerbsverläufe, und der Kohorteneffekt zu Ungunsten jüngerer Personen ist nicht oder in geringerem Umfang zu erwarten.

#### 5.2.4 Annahmen III: Folgen für Sequenzeigenschaften, inhaltliche Auswertungen und Edition

##### *Folgen für die Sequenzeigenschaften*

Was die Inzidenz bzw. Frequenz von Episoden und Übergängen betrifft, so erwarte ich, dass sich durch die Neigung des autobiografischen Erinnerns zur Komplexitätsreduktion die Anzahl, das heißt Inzidenz und Frequenz, aller Episoden und Übergänge reduziert. Nichterwerbstätigkeiten sind davon vermutlich stärker betroffen als Erwerbstätigkeiten, wohingegen die Anzahl der Arbeitslosigkeiten auch stabil geblieben oder zugenommen hat, wenn Zeiträume in der zweiten Befragung bereitwilliger als Arbeitslosigkeit beschrieben werden. Von den Übergängen werden sich besonders die Anzahl der innerbetrieblichen Wechsel reduzieren.

Die Folgen für die Dauerverteilung, also die Verteilung des betrachteten Zeitausschnitts auf die drei Episodentypen Erwerbstätigkeit, Arbeitslosigkeit und Nichterwerbstätigkeit, sind schwerer vorherzusagen. Wenn Episoden unterschlagen werden, müssen die frei gewordenen Zeiträume durch das Einführen anderer oder das Strecken der angrenzenden Episoden gefüllt werden. Hier ist anzunehmen, dass vor allem Erwerbstätigkeiten als Standardaktivität gestreckt oder an Stelle der unterschlagenen Episoden hinzugefügt werden, sodass es zu einer Zunahme der Dauer in Erwerbstätigkeit kommt. Aufgrund der besonderen Umstände des Transformationszeitraums könnte es hier auch mit der Stabilität oder Zunahme von Arbeitslosigkeitsepisoden zu einer Stabilität oder Zunahme der Dauer in Arbeitslosigkeit kommen. Zu prüfen ist hier auch die Frage, ob die Zu- bzw. Abnahme von Arbeitslosigkeiten und Nichterwerbstätigkeiten vor allem durch Erwerbstätigkeit ausgeglichen wird, oder ob es auch

Abbildung 5.4: Hypothesen im Überblick

---

A	WAS wird abweichend rekonstruiert?	
	a1 <i>Episoden</i>	Arbeitslosigkeit und Nichterwerbstätigkeit: weniger stabil
	a2 <i>Übergänge</i>	Wechsel zwischen Erwerbstätigkeit und Arbeitslosigkeit bzw. Nichterwerbstätigkeit sowie zwischenbetriebliche Wechsel: stabiler Übergänge zwischen Arbeitslosigkeit und Nichterwerbstätigkeit und innerbetriebliche Wechsel: weniger stabil
B	WER rekonstruiert abweichend?	
	b1 <i>Viele Episoden</i> bzw. Übergänge	Weniger stabil
	b2 <i>Komplexe Sequenzen</i>	Weniger stabil
	b3 <i>Geschlecht</i>	Keine Unterschiede
	b4 <i>Kohorten</i>	Keine Unterschiede
C	WIE fallen die abweichenden Rekonstruktionen aus und welche Folgen hat das?	
	c1 <i>Inzidenz bzw. Frequenz</i>	Alle Episoden und Übergänge reduziert
	c2 <i>Zusammensetzung</i>	Weniger wechselhafte und komplexe Sequenzen
	c3 <i>Dauerverteilung</i>	Erwerbstätigkeit: Zunahme Nichterwerbstätigkeit: Reduktion Arbeitslosigkeit: Stabilität oder Zunahme Erwerbstätigkeiten kompensieren für Arbeitslosigkeiten und Nichterwerbstätigkeiten (Strecken der Standardaktivität) Arbeitslosigkeit und Nichterwerbstätigkeit kompensieren für einander (Umdeuten)
	c4 <i>Timing erster Übergang</i>	Groß und nach später verschoben, wenn Fehler bei Schritt 2 Gering und symmetrisch, wenn Fehler bei Schritt 3
	c5 <i>Deskriptive und explanatorische Analysen</i>	Abweichend, wenn es um längsschnittliche und temporale Informationen (Dauer, Timing usw.) geht
D	Welche Verbesserung bringt die Edition?	
	d1 <i>Steigert die Übereinstimmung erheblich</i>	
	d2 <i>Steigert die Übereinstimmung weniger</i> für innerbetriebliche Wechsel und Umdeutungen von Arbeitslosigkeit und Nichterwerbstätigkeit und umgekehrt	

---

in einem nennenswerten Umfang dazu kommt, dass sich die beiden Lückenzustände durch Umdefinition kompensatorisch verändern.

Das *Timing* des ersten Übergangs kann aus zwei Gründen abweichend ausfallen: durch abweichende Rekonstruktion des Übergangs (Schritt 2) oder abweichende Datierung des übereinstimmend rekonstruierten Übergangs (Schritt 3). Ich nehme an, dass Fehler beim Datieren eher gering ausfallen und dass das Datum in der zweiten Befragung in gleichem Maße nach früher oder später verschoben wird. Wenn hingegen schon bei Schritt 2 der erste Übergang in der Sequenz abweichend rekonstruiert wird, sind größere Abweichungen zu erwarten. In diesem Fall erwarte ich außerdem, dass das Timing des ersten Übergangs häufiger nach hinten verschoben wird, da Übergänge häufiger vergessen als hinzugefügt werden und somit ein abweichend rekonstruierter Übergang meistens einer ist, der nach dem ersten unterschlagenen stattfand.

Die *Folgen für deskriptive und explanatorische Analysen* sind nicht leicht abzuschätzen. Wenn die verglichenen Informationen symmetrisch und nicht unabhängig vom tatsächlichen Sachverhalt abweichend rekonstruiert wurden, können die Ergebnisse trotz verbreiteter Abweichungen übereinstimmen. Ich erwarte aber, dass Analysen, die Zeitaspekte wie das genaue Timing eines Übergangs berücksichtigen, stärker voneinander abweichen, da es sich dabei um eine besonders labile Informationsdimension handelt (Paull, 2002; Pierret, 2001).

#### *Welche Verbesserung bringt die Edition?*

Für den Vergleich verwende ich die uneditierten Daten, wie sie im Anschluss an Datenerhebung und -eingabe vorlagen, denn sie stellen die originären Rekonstruktionen der Befragten dar. Ich nehme an, dass die Übereinstimmung für die edierten Daten in jeder Hinsicht höher ausfällt. Allerdings können unterschiedliche Inhalte in unterschiedlichem Maß von der Edition profitieren, denn abweichende Rekonstruktionen werden vor allem dort entdeckt, wo sie zu Widersprüchen innerhalb der Erwerbsbiografie oder mit den Angaben zu anderen Bereichen führen. Deshalb erwarte ich, dass vor allem die Episoden und Übergänge von der Edition profitieren, die auch „Spuren“ hinterlassen, wenn sie unterschlagen bzw. hinzugefügt werden, also wiederum die Übergänge zwischen Erwerbstätigkeiten einerseits und Arbeitslosigkeit und Nichterwerbstätigkeit andererseits sowie zwischenbetriebliche Wechsel.

### **5.3 Auswertung**

#### 5.3.1 Überblick über die Daten

Von 1.049 Befragten liegen vergleichbare Angaben über das doppelt abgedeckte Zeitfenster vor. Davon sind 47 Prozent Männer und ungefähr je ein Drittel gehören zu jeder der drei Kohorten (1939–41: 34%, 1949–51: 31,9% und 1959–61: 34%). Der doppelt abgedeckte Zeitraum umfasst zwischen 9<sup>4</sup> und 34 Monaten, im Durchschnitt 26 Monate. Die zweiten Befragung fand zwischen 43 und 77 Monaten nach der ersten statt; im Mittel betrug die Zunahme retrospektiver Distanz zwischen den beiden Befragungen 53,4 Monate, also etwa viereinhalb Jahre. Erwerbstätigkeiten machen in beiden Befragungen über 80 Prozent aller berichteten Episoden aus und in beiden Befragungen enthalten etwa 70 Prozent der Sequenzen nur Erwerbstätigkeiten. Frauen berichten in beiden Befragungen häufiger Sequenzen mit Arbeitslosigkeiten und Nichterwerbstätigkeiten; dies gilt auch für die jüngere Kohorte.

In der ersten Befragung werden im Zeitfenster 1.883 Episoden berichtet, in der zweiten 1.731 Episoden (Tab. 5.1, oberer Teil). Die Anzahl der Erwerbs- und Nichterwerbstätigkeiten hat in der zweiten Befragung abgenommen (wobei Nichterwerbstätigkeiten prozentual stärker betroffen sind), Anzahl und Anteil der Arbeitslosigkeiten hingegen sind um nahezu ein Viertel gestiegen. Entsprechend hat auch, trotz einer insgesamt Reduktion von Übergängen, die Anzahl der Übergänge aus oder in Arbeitslosigkeit zugenommen (Tab. 5.1, unterer Teil).

<sup>4</sup> Dies bedeutet nicht, dass die erste Befragung schon neun Monate nach dem Mauerfall stattgefunden hatte. Für sechs Fälle lagen aber, vermutlich aufgrund von Eingabe- und Aufzeichnungsfehlern, in einer der beiden Befragungen Angaben nur über einen Teil des Vergleichszeitraums vor.



Tabelle 5.1: Episoden und Übergänge LV-Ost – erste und zweite Befragung

	Erste Befragung		Zweite Befragung		Veränderung in %
	absolut	in %	absolut	in %	
Anzahl Episoden	1.883	100,0	1.731	100,0	-8,07
Erwerbstätigkeiten	1.586	84,2	1.424	82,3	-10,2
Arbeitslosigkeiten	141	7,5	176	10,2	24,8
Nichterwerbstätigkeiten	156	8,3	131	7,6	-16,0
Anzahl Übergänge	834	100,0	682	100,0	-18,2
ET-ET innerbetrieblich	159	19,6	129	18,9	-18,9
ET-ET zwischenbetrieblich	329	39,5	206	30,2	-37,4
ET-ALO	125	14,9	145	21,3	16,0
ET-NET	80	9,6	59	8,7	-26,3
ALO-ET	52	6,2	66	9,7	26,9
ALO-NET	8	1,0	18	2,6	125,0
NET-ET	70	8,4	44	6,5	-37,1
NET-ALO	11	1,3	15	2,2	36,4

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

Übergänge zwischen Erwerbstätigkeiten und zwischen Erwerbstätigkeit und Nichterwerbstätigkeit haben hingegen abgenommen, wobei wider Erwarten zwischenbetriebliche Übergänge stärker betroffen sind als innerbetriebliche.

Der Vergleich der Information auf Stichprobenebene bestätigt die Annahmen bereits teilweise. Besonders die Zunahme von Arbeitslosigkeiten, die im Widerspruch zu der bisherigen Befundlage steht, deutet darauf hin, dass die Transformation spezielle Bedingungen für die Rekonstruktion von Arbeitslosigkeit bietet. Vergleiche auf Stichprobenebenen sind allerdings nur bedingt aussagekräftig, da die Veränderungen immer Nettoeffekte von individuellen „Wanderbewegungen“ sind. Im nächsten Abschnitt werde ich deshalb nun auf der Individual-ebene den tatsächlichen Fehler bei der Rekonstruktion der Episoden- und Dauerinformation der Erwerbsverläufe betrachten.

### 5.3.2 Qualitative Einzelfallbetrachtungen

Zuerst werde ich eine qualitative Einzelfallbetrachtung durchführen, um ein anschauliches Verständnis dafür zu vermitteln, wie die Sequenzen „umstrukturiert“ werden. An einer Stichprobe von 150 Erwerbsverläufen, die in den beiden Datensätzen eine deutlich unterschiedliche Anzahl von Episoden oder eine unterschiedliche Dauerverteilung aufwiesen, habe ich elf Arten der Abweichung identifiziert, die mit dem dreistufigen Rekonstruktionsprozess in Verbindung gebracht werden können<sup>5</sup>. Diese sind in Abbildung 5.5 anhand ausgewählter Einzelfälle veranschaulicht.

<sup>5</sup> Freilich verbleiben Abweichungen, die keinem der Phänomene eindeutig zugeordnet werden können. Ich gehe aber davon aus, dass sich auch diese Fälle prinzipiell bei besserem Verständnis der individuellen und kontextuellen Hintergründe zuordnen ließen.

Abbildung 5.5: Elf Arten gedächtnisversuchter Abweichung

		1990												1991												1992																																																											
		12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9																																																		
1	Strauchen und Strecken																																																																																				
erste	ET 1																									ALO																																																											
zweite	ET 1																									ALO																																																											
2a	Auslassen und Strecken																																																																																				
erste	ET 1													NET												ET 2																																																											
zweite	ET 1													ET 2																																																																							
2b	Auslassen und Strecken																																																																																				
erste	ET 1													ALO 1												ET 2																																																											
zweite	ET 1																									ET 3																																																											
2c	Auslassen und Strecken																																																																																				
erste	ET 1													ALO 1												ET 2												ALO 2												ET 3												ALO 3												ET 4											
zweite	ET 1													ET 2												ALO 1																																				ET 4																							
3	Einbauen und Stauchen																																																																																				
erste	ET 1													ET 2																																																																							
zweite	ET 1	ALO																								ET 2																																																											
4	Nach vorne aus Zeitfenster hinausdatieren																																																																																				
erste	ET 1 Molkerei													ET 2 Kurhotel												ALO																																																											
zweite	ET 1 Hotel																									ALO																																																											

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

noch Abbildung 5.5: Elf Arten gedächtnisverursachter Abweichung

		1990												1991												1992											
		12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7				
5	Von vorne in Zeitfenster hineindatieren																																				
erste	ALO																																				
zweite	ET																																				
	ALO																																				
6	Nach hinten aus Zeitfenster hinausdatieren																																				
erste	ET 1   ALO																																				
	ET 2																																				
zweite	NET																																				
7	Von hinten in Zeitfenster hineindatieren																																				
erste	ET 1   ET 2																																				
	ET 3																																				
zweite	ET																																				
	NET																																				
8a	Episode umdefinieren																																				
erste	ET 1																																				
	ET 2 (Kurzarbeit Null)																																				
zweite	ET 1																																				
	ALO																																				
8b	Episode umdefinieren																																				
erste	NET (Hausfrau)																																				
zweite	ALO																																				
	ET																																				
9a	Übergang umdefinieren (von zwischen- zu innerbetrieblich)																																				
erste	ET 1 (Ingenieur)																																				
	Wechsel zwischenbetrieblich																																				
zweite	ET 1 (Elektroingenieur)																																				
	Wechsel innerbetrieblich																																				
	ET 2 (ABM)																																				
	ET 2 (Transportabteilung)																																				

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

noch Abbildung 5.5: Elf Arten gedächtnisversuchter Abweichung

		1990												1991												1992											
		1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7					
9b	Übergang umdefinieren (von inner- zu zwischenbetrieblich)																																				
erste	ET 1 (Oberstufenlehrer)													Wechsel innerbetrieblich												ET 2 (Gymnasiallehrer)											
zweite	ET 1 (Lehrer POS)													Wechsel zwischenbetrieblich												ET 2 (Gymnasiallehrer)											
10a	Fusion ähnlicher Episoden/Vergessen von Übergängen																																				
erste	ET 1 (VEB X)													ET 2 (Kraftfahrer Amt X)																							
zweite	ET (VEB X GmbH)																									ET 2											
10b	Fusion ähnlicher Episoden/Vergessen von Übergängen																																				
erste	ET 1 (Werksratmeister)																									ET 2 (ABM Firma X)											
zweite	ET (Meister Firma X)																																				
11a	Defusion von Episoden/Einbauen neuer Übergänge																																				
erste	ET 1 (Verkäuferin Chemiefaser)																																				
zweite	ET 1 (Verkäuferin Chemiefaser)																									ET 2 (Kurzarbeit Null)											
11b	Defusion von Episoden/Einbauen neuer Übergänge																																				
erste	NET (Hausfrau)																																				
zweite	NET (Hausfrau)																									ET (Reinigungskraft, 15 Std./Woche, 450,- DM/Monat)											
11c	Defusion von Episoden/Einbauen neuer Übergänge																																				
erste	ET 1 (HO Waren d. tägl. Bedarfs, Verkäufer)													ET 2 (Filialleiter Supermarkt X)																							
zweite	ET 1 (HO, stellv. Filialleiter)													ET 2 (Lebensmittel Supermarkt Y)												ET 3 (Handelskette X, stellv. Filial.) ET 4 (stellv. Filialleiter)											
erste	Keine plausible Erklärung																																				
zweite	ET 1													NET (Erziehungsurlaub)																							
	NET (Babyjahr)													ALO																							

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

*Fall 1: Stauchen und Strecken*

Die Befragungsperson 1 hat bei Schritt 1 und 2 identische Episoden und Übergänge rekonstruiert, bei Schritt 3 aber das Ende der ersten Episode in der zweiten Befragung um einen Monat nach hinten datiert bzw. die erste Episode etwas gestreckt und die zweite etwas gestaucht. Dadurch werden die Dauerverteilung und das Timing des ersten Übergangs verändert. Hätte es sich bei der zweiten Episode ebenfalls um eine Erwerbstätigkeit gehandelt, wäre nur das Timing betroffen gewesen. Praktisch ist der Fehler in diesem Fall unbedeutend.

*Fall 2: Auslassen und Strecken*

In den Fällen 2a, 2b und 2c haben die Befragten bei Schritt 2 einen Übergang nicht rekonstruiert, dadurch eine in der ersten Befragung angegebene Episode komplett unterschlagen und den freigewordenen Zeitraum durch das Strecken einer angrenzenden Episode aufgefüllt. Im Falle 2a wird nur ein Monat Hausfrauentätigkeit eingeebnet. Im Fall 2b hingegen werden schon fünf Monate, im Fall 2c sieben Monate Arbeitslosigkeit unterschlagen. Dadurch werden ihre Lebensverläufe aus der längeren retrospektiven Distanz heraus auf Kosten der weniger verbreiteten Nichterwerbstätigkeiten und Arbeitslosigkeiten vereinfacht, geglättet und konventionalisiert. Inzidenz und Frequenz der unterschlagenen Episoden sinken, und mit ihr verändern sich Dauerverteilungen und teilweise das Timing der Übergänge.

*Fall 3: Einbauen und Stauchen*

Der Befragte rekonstruiert bei Schritt 2 einen Übergang in einem Zeitraum, den er in der ersten Befragung durchgehend als Erwerbstätigkeit beschrieben hatte, und beschreibt einen Teil des Zeitraums als Arbeitslosigkeit. Für die hinzugefügte Arbeitslosigkeit wird die Erwerbstätigkeit sozusagen gestaucht. Als Folge der abweichenden Rekonstruktion reduzieren sich die Monate in Erwerbstätigkeit, die Monate in Arbeitslosigkeit hingegen steigen und der erste Übergang findet früher statt. Auch hier bleiben zumindest Inzidenz und Reihenfolge der Erwerbstätigkeiten unverändert.

*Fälle 4–7: Zeitfensterphänomene: Hinaus- bzw. Hineindatieren*

Prinzipiell handelt es sich um Strecken und Stauchen bei übereinstimmender Rekonstruktion von Episoden und Übergängen wie in Fall 1. Dabei wird aber das Ende bzw. der Anfang einer Episode nach vorne oder hinten aus dem Zeitfenster hinaus oder in das Zeitfenster hinein verlegt, wodurch sich Inzidenzen und Frequenzen verändern. Ob ein Zeitfensterphänomen vorliegt oder ob eine Episode unterschlagen bzw. neu eingefügt wurde (Fall 2 und 3), kann praktisch nur durch Information über die Zeiträume nach oder vor dem Zeitfenster getätigt werden, wenn sich hier die Episoden bzw. Übergänge finden, die in der jeweils anderen Befragung innerhalb des Zeitfensters auftauchen.

*Fall 8: Episode umdefinieren*

Im Fall 8a rekonstruiert der Befragte bei Schritt 2 übereinstimmend, dass und wann Veränderungen in zentralen erwerbs- und arbeitsmarktbezogenen Rollen und Aktivitäten stattfanden. Allerdings bezeichnet er den Lebensabschnitt in Schritt 1 nicht mehr als Erwerbstätigkeit, sondern als Arbeitslosigkeit. In einem Freitextfeld ist vermerkt, dass er in dem betreffenden Zeit-

raum auf Kurzarbeit Null gesetzt war. Der Wechsel zu Kurzarbeit bringt also Veränderungen mit sich, die einem Übergang in Arbeitslosigkeit ähneln; Kurzarbeit Null ist im Hinblick auf die angebotenen Zustände ambivalent und damit potenziell instabil repräsentiert. Dieser Befragte hat diese Ambivalenz in der zweiten Befragung anders aufgelöst als in der ersten Befragung.

In Fall 8b beschreibt die Befragte bei Schritt 1 einen Zustand nicht mehr als Hausfrau, sondern als Arbeitslosigkeit (und streckt außerdem noch nach vorne bis zum vorderen Ende des Zeitfensters und nach hinten bis über den Interviewzeitpunkt hinaus). Ihre erwerbs- und arbeitsmarktbezogenen Rollen und Aktivitäten in diesem Zeitraum erlauben sowohl die Beschreibung als Arbeitslosigkeit als auch als Nichterwerbstätigkeit und die Zweideutigkeit wurde in der zweiten Befragung anders aufgelöst als in der ersten. Von solchen Fehlern wird die gesamte Episodenstruktur affektiert, Inzidenz und Dauer verändern sich, die Komplexität nimmt zu, andere Übergänge finden statt; in Fall 8a bleibt aber das Timing des ersten Übergangs weitgehend stabil.

#### *Fall 9: Übergänge umdefinieren*

Fall 9a rekonstruiert Episoden und Übergänge bei Schritt 2 und 3 übereinstimmend, beschreibt aber den Übergang in der ersten Befragung als zwischenbetrieblichen und in der zweiten Befragung als innerbetrieblichen Wechsel. Da es sich bei der zweiten Erwerbstätigkeit um eine Arbeitsbeschaffungsmaßnahme des Arbeitsamts handelt, ist hier der Begriff des „Betriebs“ weniger eindeutig. Fall 9b, ein Lehrer, wechselt umgekehrt von inner- zu zwischenbetrieblichem Wechsel. Für Angestellte im öffentlichen Dienst und Beamte ist es ohnehin schwieriger, Stellenwechsel voneinander abzugrenzen, wie Becker (1993) und de Graaf und Wegener (1989) gezeigt haben, und umso mehr in einer Situation, in der der öffentliche Dienst als institutionelles Umfeld für Berufsverläufe völlig neu eingeführt wurde. Von Fehlern dieser Art werden Inzidenz, Zusammensetzung, Dauerverteilungen und Datum des Übergangs erfreulich wenig beeinträchtigt.

#### *Fall 10: Fusion ähnlicher Episoden/Vergessen von Übergängen*

Werden bei Schritt 2 Übergänge zwischen ähnlichen Episoden nicht rekonstruiert, können zwei oder mehr Episoden eines Typs zu einer Episode „fusioniert“ werden. Fall 10a ebnet in der zweiten Befragung einen Wechsel vom „VEB Berlin“ zum „Garten- und Gemüseladeamt Pan-kow“ ein. Es ist anzunehmen, dass es sich hierbei vor allem um einen Wechsel der Rechtsform des Betriebs von VEB zu GmbH handelt, während die Tätigkeiten, Identität und sozialen Rollen des Befragten sich nicht nachhaltig geändert haben, sodass kein neuer Lebensabschnitt repräsentiert ist. Ähnlich rekonstruiert der Befragte 10b den Wechsel vom Werkstattmeister zur ABM-Kraft in der zweiten Befragung nicht mehr und berichtet eine durchgehende Erwerbsepisode. Der Effekt auf die Sequenz ist unterschiedlich: Zwar verringern sich die Episodenzahl und die Anzahl der Übergänge, aber die Dauer in Erwerbstätigkeit bleibt unverändert<sup>6</sup>.

<sup>6</sup> Da Arbeitslosigkeiten keine inneren Unterteilungen haben und aneinander grenzende Nichterwerbstätigkeiten bereits für den Vergleich zusammengelegt wurden, kommen Fusion und Defusion nur bei Erwerbstätigkeiten vor. Wenn Episoden unterschiedlichen Typs fusioniert werden – etwa eine geringfügige Beschäftigung und eine Arbeitslosigkeit –, muss der Zeitraum einheitlich beschrieben werden, weswegen dann eine Umdeutung vorliegt (siehe Fall 8).

*Fall 11: Defusion von Episoden /Einbauen eines neuen Übergangs*

Hierbei handelt es sich um den komplementären Vorgang zur Episodenfusion: Lebensabschnitte, in denen aus der kürzeren retrospektiven Distanz kein Wechsel wahrgenommen wurde, werden aus der längeren Rückschau als zwei oder mehrere unterscheidbare Abschnitte rekonstruiert. Person 11a hat in der ersten Befragung bei Schritt 2 keine Veränderung berichtet, rekonstruiert aber in der zweiten Befragung einen Übergang in Kurzarbeit Null, der im Jahre 1995, nach dem Ende des Zeitfensters, eine Arbeitslosigkeit gefolgt. Diese Person, die in der ersten Befragung vermutlich noch mit der Weiterführung ihrer Erwerbstätigkeit rechnete, bringt also die Rekonstruktion der Vergangenheit in Einklang mit ihren jeweils aktuellen Erfahrungen. Person 11b rekonstruiert nur in der zweiten Befragung die Aufnahme einer geringfügigen Beschäftigung als Reinigungskraft, welche ebenfalls 1995 von einer Vollzeitbeschäftigung gefolgt wird. In der ersten Befragung erschien die Erwerbstätigkeit vermutlich der „eigentlichen“ Hausfrauentätigkeit untergeordnet, während in der längeren Rückschau diese erste geringfügige Beschäftigung als Wiedereinstieg eingeschätzt und deshalb auch als eigenständiger Lebensabschnitt rekonstruiert wird. Fall 11c zeigt einen Befragten, bei dem deutlich Defusionsphänomene auftreten, ohne dass man genau sagen könnte, warum gerade in dieser Weise defusioniert wird und was für ein tatsächlicher Verlauf diesen beiden abweichenden Wahrnehmungen zu Grunde liegt.

Die qualitative Fehlerbetrachtung hat gezeigt, dass in der längeren Rückschau mitunter tief greifende Umstrukturierungen vorgenommen werden, die mit den drei kognitiven Schritten der Datengenerierung und der Repräsentation der Erwerbsgeschichte als Lebensabschnitte in Verbindung gebracht werden können. Die Veränderungen können sowohl auf eine Zu- als auch eine Abnahme an Komplexität, Wechselhaftigkeit und Konventionalität hinauslaufen. Die Folgen für die ausgewählten Informationsdimensionen sind teilweise gravierend und teilweise unbedeutend.

Es lässt sich allerdings keine direkte Verbindung zwischen bestimmten Arten der Umstrukturierung und bestimmten Veränderungen von Eigenschaften der Sequenz herstellen. Zum einen können dieselben Veränderungen an der Gesamtsequenz als Resultat unterschiedlicher Umstrukturierungen auftreten. Beispielsweise reduzieren sich Episoden durch Unterschlagen (Fall 2a), Umdeuten (Fall 8a) und Fusion von Episoden (Fall 10b). Zum anderen können an einer Sequenz mehrere Umstrukturierungen vorgenommen werden, die additiv und interaktiv die Gesamtveränderung an der Sequenz herstellen wie in Fall 2b, wo zusätzlich zur unterschlagenen Arbeitslosigkeit an anderer Stelle eine Erwerbstätigkeit defusioniert wird. Da sich die Veränderungen der Sequenzeigenschaften „netto“ durch die Wirkung aller Einzelveränderungen ergeben, ist auch umgekehrt kein eindeutiger Rückschluss auf die zu Grunde liegenden Umstrukturierungen möglich.

### 5.3.3 Folgen für Inzidenz und Frequenz von Episoden

*Die Episodenstruktur: Inzidenz und Frequenz von Episoden und Übergängen*

Tabelle 5.2 zeigt für die 1.025 Vergleichspersonen den Fehler für die Abbildung der Sequenzinformation Inzidenz und Frequenz der drei Episodentypen. Dabei sind immer nur die Befragten berücksichtigt worden, die in mindestens einer der beiden Befragungen eine Episode des jeweiligen Typs angegeben hatten.

Tabelle 5.2: Fehler bei Inzidenz/Frequenz der Episoden

	Erwerbstätigkeiten				Arbeitslosigkeiten				Nichterwerbstätigkeiten			
	N	Irrtums- risiko	Fehler- größe	Fehler- rich- tung	N	Irrtums- risiko	Fehler- größe	Fehler- rich- tung	N	Irrtums- risiko	Fehler- größe	Fehler- rich- tung
Alle	1.025	31,5	0,36	72,2	195	48,7	0,50	31,6	195	62,1	0,32	58,7
Männer	490	32,8	0,36	79,2	76	50,0	0,55	36,8	59	72,9	0,22	39,5
Frauen	535	30,3	0,35	70,2	119	48,9	0,48	28,1	136	57,3	0,37	69,2
Kohorte um												
1940	348	25,2	0,30	68,9	66	55,0	0,58	21,2	53	67,9	0,35	63,9
1950	325	33,5	0,39	69,7	59	45,8	0,47	40,7	57	57,9	0,32	51,5
1960	352	35,2	0,39	74,7	76	46,1	0,47	34,3	85	61,2	0,29	59,6

Irrtumsrisiko: Prozentualer Anteil der abweichenden an allen Angabenpaaren.

Fehlergröße: Mittlere numerische Abweichung ohne Vorzeichen.

Fehlerrichtung: Prozentualer Anteil der abweichenden Angaben, bei denen in der zweiten Befragung eine geringere Anzahl von Episoden angegeben wird.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

Das Risiko, eine abweichende Anzahl von Erwerbstätigkeiten anzugeben, liegt bei 31,5 Prozent. Das heißt, dass ein knappes Drittel der Befragten unterschiedlich viele Erwerbsepisoden angibt. Für die Anzahl der Arbeitslosigkeiten liegt das Irrtumsrisiko bei 48,7 Prozent und für die Anzahl der Nichterwerbstätigkeiten bei 62,1 Prozent. Das Ausmaß der Abweichungen (die durchschnittliche numerische Differenz der Episodenanzahlen) ist gering; am häufigsten wird nur eine Episode weniger oder mehr berichtet.

Die *Richtung* des Fehlers ist dargestellt als Anteil derjenigen, die in der zweiten Befragung eine geringere Anzahl von Episoden an allen abweichenden Angabenpaaren angeben. Für Erwerbstätigkeiten berichteten 72,2 Prozent der Personen eine geringere Anzahl, für Nichterwerbstätigkeiten sind dies 58,7 Prozent und für Arbeitslosigkeiten nur 31,6 Prozent. Die insgesamt Neigung zur Reduktion für Erwerbs- und Nichterwerbstätigkeiten kommt also dadurch zu Stande, dass eine größere Anzahl Befragter die Anzahl eher reduzieren als erhöhen, wobei bei Nichterwerbstätigkeiten beide Richtungen ausgewogener sind. Bei Arbeitslosigkeit hingegen überwiegen die Personen, die mehr Episoden berichten, und verursachen damit eine allgemeine Zunahme von Arbeitslosigkeitsepisoden.

In Tabelle 5.2 sind Irrtumsrisiko, Fehlergröße und -richtung für Geschlechter und Kohorten getrennt aufgelistet. Frauen haben für alle Episodentypen wie erwartet kein höheres, sondern sogar ein etwas geringeres Irrtumsrisiko, besonders bei Nichterwerbstätigkeiten. Außerdem haben Männer und Frauen bei Nichterwerbstätigkeiten umgekehrte Fehlerrichtungen: Während Frauen Nichterwerbstätigkeiten häufiger unterschlagen als hinzufügen, ist es für Männer umgekehrt der Fall, dass sie Nichterwerbstätigkeiten häufiger hinzufügen. Dies ist genau der umgekehrte Trend, wie ihn Paull (2002) für britische Frauen nachgewiesen hatte. Diese hatten häufiger Nichterwerbstätigkeiten als Männer hinzugefügt, indem sie Arbeitslosigkeiten in Nichterwerbstätigkeiten umdefinierten.

Kohortenunterschiede treten ebenfalls auf, folgen aber keinem klaren Muster. Die älteste Kohorte der um 1940 Geborenen macht bei der Anzahl der Erwerbstätigkeiten am seltensten Fehler, bei Arbeitslosigkeiten und Nichterwerbstätigkeiten hingegen am häufigsten. Sie neigt



Tabelle 5.3: Inzidenz/Frequenz der Episoden – Veränderungen zwischen erster und zweiter Befragung

		0		1		2		3 bis 5		Gesamt
		absolut	in %	absolut	in %	absolut	in %	absolut	in %	
Anzahl ET zweite Befragung										
Anzahl ET erste Befragung	0	24	80,0	6	20,0					30
	1	8	1,4	481	86,4	61	10,9	7	1,3	557
	2	7	1,9	152	41,0	197	53,1	15	4,0	371
	3–5	0		21	23,7	38	41,8	32	35,2	91
	Gesamt	39		660		296		54		1.049
Anzahl ALO zweite Befragung										
Anzahl ALO erste Befragung	0	854	93,2	60	6,6	2	0,2			916
	1	26	20,6	97	76,9	3	2,4			126
	2–3	1		3		3				7
	Gesamt	881		160		8				1.049
Anzahl NET zweite Befragung										
Anzahl NET erste Befragung	0	854	94,8	45	5,0	1	0,1			900
	1	64	45,1	74	52,1	4	2,8			142
	2	5		2						7
	Gesamt	923		121		5				1.049

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

außerdem am stärksten dazu, Nichterwerbstätigkeiten zu unterschlagen und Arbeitslosigkeiten hinzuzufügen. Die jüngste Kohorte macht am häufigsten Fehler bei der Anzahl der Erwerbsepisoden, ist bei Arbeitslosigkeiten und Nichterwerbstätigkeiten hingegen nicht unterlegen. Bei Nichterwerbstätigkeiten ist es die mittlere Kohorte, die das geringste Irrtumrisiko aufweist.

Die individuellen „Wanderbewegungen“ werden in Tabelle 5.3 detaillierter dargestellt. Für Erwerbstätigkeiten finden Wanderbewegungen zwischen allen Zellen statt, das heißt, es gibt Personen, die bis zu drei Erwerbstätigkeiten mehr oder weniger angeben. Das Irrtumrisiko steigt deutlich mit der Anzahl der Erwerbstätigkeiten in der ersten Befragung: von 20 Prozent bei keiner und 13,6 Prozent bei nur einer Erwerbsepisode auf bereits 46,9 Prozent bei zwei Erwerbsepisoden und auf 64,8 Prozent in der vier bis acht Erwerbsepisoden. Wechselhafte Verläufe werden also deutlich instabiler erinnert, und in der Regel fällt die Sequenz in der zweiten Befragung dabei weniger wechselhaft aus. Dies kann auch als Abhängigkeit der Abweichungen vom wahren Sachverhalt begriffen werden.

Bei Arbeitslosigkeiten und Nichterwerbstätigkeiten hingegen spielen sich nahezu alle Wanderbewegungen zwischen einer und keiner Episode ab, da nur sehr wenige Teilnehmer mehr als eine solche Episode berichten. Veränderungen laufen also fast immer darauf hinaus, dass Arbeitslosigkeit bzw. Nichterwerbstätigkeit ganz aus dem Zeitfenster getilgt oder ganz

neu hinzugefügt wird. Wer in der ersten Befragung keine Arbeitslosigkeit berichtet hat, berichtet in über 90 Prozent der Fälle auch in der zweiten Befragung keine, und wer in der ersten Befragung eine berichtet hat, tut dies in 76,9 Prozent der Fälle auch in der zweiten. Ähnlich stellt sich die Sachlage für Nichterwerbstätigkeiten dar. Das Irrtumsrisiko ist für diejenigen ohne Nichterwerbstätigkeit in der ersten Befragung sehr gering, diejenigen mit einer Nichterwerbstätigkeit haben aber ein Risiko von fast 50 Prozent, diese in der zweiten Befragung zu unterschlagen.

*Zusammensetzung und Komplexität der Sequenz*

Die Erwerbssequenzen bestehen entweder nur aus Episoden eines Typs (nahezu immer handelt es sich dabei um Erwerbstätigkeiten), Episoden zweier verschiedener Typen oder aus Episoden aller drei Typen. Durch die Erinnerungsprozesse sollte sich retrospektiv die Komplexität und Wechselhaftigkeit der erinnerten Sequenzen reduzieren, und komplexere, wechselhaftere Sequenzen sollten eine höhere Instabilität aufweisen. In Tabelle 5.4 sehen wir an den Randsummen, dass in den beiden Befragungen eine nahezu identische Anzahl von Verläufen mit nur einem, zwei und allen drei Episodentypen berichtet wird. In der zweiten Befragung treten etwas häufiger Verläufe auf, die aus Erwerbstätigkeiten und Arbeitslosigkeiten bestehen, und etwas seltener solche, die aus Erwerbstätigkeiten und Nichterwerbstätigkeiten bestehen.

In den einzelnen Zellen sehen wir, dass diese Veränderungen wieder Nettoeffekte von Wanderbewegungen sind. Am stabilsten werden die einfachsten Sequenzen rekonstruiert: Diejenigen, die in der ersten Befragung nur Episoden eines einzigen Typs angegeben haben, tun dies in über 90 Prozent der Fälle in der zweiten Befragung wieder und berichten ansonsten gleich häufig eine zusätzliche Arbeitslosigkeit bzw. Nichterwerbstätigkeit. Sequenzen aus Erwerbstätigkeiten und Arbeitslosigkeiten werden in zwei Drittel der Fälle auch in der zwei-

Tabelle 5.4: Zusammensetzung und Komplexität der Erwerbssequenzen – Veränderungen zwischen erster und zweiter Befragung

Erste Befragung	Zweite Befragung										Gesamt	
	Nur Episoden eines Typs		ET und ALO		ET und NET		ALO und NET		Episoden aller drei Typen			
	absolut	in%	absolut	in%	absolut	in%	absolut	in%	absolut	in%		
Nur Episoden eines Typs	764	92,8	30	3,7	27	3,3	1		1		823	78,5
ET und ALO	18	17,8	67	66,3	5	4,9	3	3,0	8	7,9	101	9,6
ET und NET	38	39,1	19	19,6	30	30,9	3	3,1	7	7,2	97	9,2
ALO und NET							1		1		2	
Episoden aller drei Typen	1		9	32,1	5	7,5	1		10	38,5	26	2,5
Gesamt	821	78,3	125	11,9	67	6,3	9	0,9	27	2,6	1.049	
					201	19,2						

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

ten Befragung als solche beschrieben. Die anderen eliminieren in etwa der Hälfte der Fälle die Arbeitslosigkeit, ansonsten fügen sie auch häufig eine Nichterwerbstätigkeit hinzu oder deuten die Arbeitslosigkeit in eine Nichterwerbstätigkeit um. Sequenzen aus Erwerbstätigkeiten und Nichterwerbstätigkeiten werden sogar nur in ein Drittel der Befragung erneut als solche beschrieben, am häufigsten wird die Nichterwerbstätigkeit gänzlich getilgt. Die sehr komplexen Sequenzen mit drei verschiedenen Episodentypen werden wieder etwas stabiler erinnert, meistens fehlt in der zweiten Befragung die Nichterwerbstätigkeit.

Insgesamt werden also wechselhaftere Episoden weniger konsistent rekonstruiert und die Veränderungen laufen auf eine Reduktion der Komplexität hinaus. Wie diese sich genau gestaltet, hängt von der Art der Episoden ab. Dies kann auch als Abhängigkeit der Abweichungen vom wahren Sachverhalt begriffen werden.

#### 5.3.4 Folgen für die Dauerverteilung

In Tabelle 5.5 sind Irrtumsrisiko, Fehlergröße und Fehlerrichtung für die Anzahl der Monate in Erwerbstätigkeit, Arbeitslosigkeit und Nichterwerbstätigkeit dargestellt. Die Angaben wurden auch dann als übereinstimmend gewertet, wenn die Abweichung nur einen Monat betrug. Außerdem wurden diejenigen Befragten ausgeschlossen, die beide Male keine Episode des jeweiligen Typs berichten. Das Irrtumsrisiko für die Anzahl der Monate in Arbeitslosigkeit und Nichterwerbstätigkeit liegt bei über 60 Prozent, während es für Erwerbstätigkeit nur 16 Prozent beträgt. Die durchschnittliche Abweichung ist ebenfalls für Arbeitslosigkeit und Nichterwerbstätigkeit deutlich höher, am höchsten bei Arbeitslosigkeit mit einer mittleren Abweichung von 6,5 Monaten. Für Erwerbstätigkeit und Nichterwerbstätigkeit ist eine Reduktion der Monate etwas häufiger, für Arbeitslosigkeit hingegen eine Zunahme deutlich häufiger.

Tabelle 5.5: Fehler bei der Dauerverteilung der drei Episodentypen

	Monate in ET				Monate in ALO				Monate in NET			
	<i>N</i>	Irrtums- risiko	Fehler- größe	Fehler- rich- tung	<i>N</i>	Irrtums- risiko	Fehler- größe	Fehler- rich- tung	<i>N</i>	Irrtums- risiko	Fehler- größe	Fehler- rich- tung
Alle	1.025	16,8	1,54	54,7	195	64,1	6,51	28,2	195	66,2	5,55	59,7
Männer	490	12,9	0,98	52,4	76	61,8	4,06	35,3	59	74,6	4,92	45,5
Frauen	535	20,1	2,06	55,9	119	65,5	7,31	24,6	136	62,5	5,03	67,1
Kohorte um												
1940	348	13,5	1,43	61,7	60	68,3	6,90	21,9	53	62,3	5,21	69,7
1950	325	11,1	1,32	63,6	59	67,8	5,24	37,5	57	54,4	5,14	48,4
1960	352	21,0	2,04	45,7	76	57,9	5,97	36,2	85	77,5	6,05	60,0

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Irrtumsrisiko: Prozentualer Anteil der abweichenden an allen Angabenpaaren.

Fehlergröße: Mittlere numerische Abweichung ohne Vorzeichen.

Fehlerrichtung: Prozentualer Anteil der abweichenden Angaben, bei denen in der zweiten Befragung eine kürzere Dauer im jeweiligen Zustand angegeben wird.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

Bei der Dauerverteilung folgen die Geschlechterunterschiede wieder teilweise dem aus anderen Studien bekannten Muster: Frauen irren sich bei Erwerbstätigkeit und Arbeitslosigkeit häufiger und machen größere Fehler als Männer, bei Nichterwerbstätigkeit irren sie sich seltener und machen etwa gleich große Fehler. Auch hier neigen Frauen dazu, die Zeit in Nichterwerbstätigkeit in der zweiten Befragung zu reduzieren, was für Männer nicht der Fall ist. Kohortenunterschiede sind weiterhin uneindeutig; die Mitglieder der jüngsten Kohorte machen über die Anzahl der Monate in Erwerbstätigkeit und Nichterwerbstätigkeit am häufigsten inkonsistente Angaben und haben größere Abweichungen als die beiden älteren Kohorten, bei Arbeitslosigkeit ist es genau umgekehrt.

Da die Erhebung ein festgelegtes Zeitfenster abdeckt, sind Abweichungen der Sequenzeigenschaften Inzidenz und Frequenz einerseits sowie Dauer andererseits interdependent: Für jede hinzugefügte oder unterschlagene Episode muss an anderer Stelle kompensiert werden, entweder durch das Hinzufügen oder Unterschlagen oder durch das Strecken oder Stauchen anderer Episoden. Im nächsten Absatz werde ich diese Interdependenz der Fehler auf drei Arten näher beleuchten. Zuerst werde ich überprüfen, wie viele der Sequenzen eine äquivalente Episodenstruktur aufweisen. Bei diesen äquivalent rekonstruierten Sequenzen erwarte ich auch eine höhere Übereinstimmung für die Dauerverteilungen. Danach werde ich überprüfen, welche Folgen die Veränderung der Anzahl der Episoden eines Typs für die Zeit hat, die auf diesen Episodentyp entfällt. Und zuletzt werde ich die Episodenstruktur der Sequenz nicht mehr berücksichtigen und überprüfen, auf welche Weise sich bei der Veränderung der Dauer in einem Episodentyp die Dauer in anderen Episodentypen kompensatorisch verändert.

### 5.3.5 Interdependenz von Episoden und Dauerverteilungen

#### *Äquivalenz der Episodenstruktur*

Ich betrachte Sequenzen als äquivalent rekonstruiert, wenn sie aus derselben Anzahl von Erwerbstätigkeiten, Arbeitslosigkeiten und Nichterwerbstätigkeiten bestehen und dieselben Übergänge aufweisen. In Tabelle 5.6 sehen wir, dass 56,1 Prozent der Teilnehmer äquivalente Erwerbssequenzen rekonstruieren. Je höher die Anzahl der Episoden in der ersten Befragung, desto geringer ist der Anteil äquivalenter Sequenzen: Bei einer Episode sind es 84,3 Prozent und bei vier oder mehr Episoden nur mehr 12,2 Prozent. Auch komplexere Sequenzen mit zwei oder mehr Episodentypen werden seltener äquivalent rekonstruiert; allerdings gibt es hier deutliche Unterschiede innerhalb der Zwei-Episoden-Sequenzen: Während Sequenzen aus Erwerbstätigkeiten und Arbeitslosigkeiten etwa zur Hälfte äquivalent sind, sind Sequenzen aus Erwerbstätigkeiten und Nichterwerbstätigkeiten nur in 16,5 Prozent der Fälle äquivalent. Zwischen Männern und Frauen bestehen keine Unterschiede, hingegen weist die älteste Kohorte 61,6 Prozent und die jüngste nur 49,6 Prozent äquivalent rekonstruierte Sequenzen auf, wobei die Erwerbsverläufe der jüngsten Kohorte mehr Episoden und Übergänge enthalten und häufiger aus zwei oder drei Episodentypen bestehen.

In Tabelle 5.6 sehen wir die Folgen nichtäquivalenter Sequenzrekonstruktion für die Dauerverteilungen. In jeder Hinsicht stimmen die Dauerinformationen deutlich besser überein, wenn die Sequenzen äquivalent rekonstruiert werden. Dies gilt besonders für Arbeitslosigkeit

Tabelle 5.6: Äquivalent rekonstruierter Sequenzen – Ursachen und Folgen

Ursachen	Geschlecht		Kohorte um				N Episoden in der ersten Befragung				Zusammensetzung in der ersten Befragung			
	alle	männlich	weiblich	1940	1950	1960	1	2	3	4+	nur Episo- den eines Typs	ET und ALO	ET und NET	ET und NET aller drei Typen
Äquivalente Sequenzen	589	275	314	220	192	177	393	155	35	6	516	49	16	7
	56,1%	55,8%	56,5%	61,6%	57,3%	49,6%	84,3%	39,0%	25,4%	12,2%	62,7%	48,5%	16,5%	26,9%
N	1.049	493	556	357	335	357	466	397	138	49	823	101	97	26

Folgen	Monate in Erwerbstätigkeit				Monate in Arbeitslosigkeit				Monate in Nichterwerbstätigkeit				N
	Irrtums- risiko in %	Fehler- größe in %	Fehler- richtung in %	Irrtums- risiko in %	Fehler- größe in %	Fehler- richtung in %	Irrtums- risiko in %	Fehler- größe in %	Fehler- richtung in %	Irrtums- risiko in %	Fehler- größe in %	Fehler- richtung in %	
Äquivalente Sequenzen	5,2	0,3	36,8	2,4	0,2	57,1	0,7	2,3	45,5	589	56,1		
Nichtäquivalente Sequenzen	43,3	3,1	56,8	24,1	2,4	29,3	27,2	2,3	45,5	460	43,9		

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung; eigene Berechnungen.

und Nichterwerbstätigkeit. Veränderungen der Episodenstruktur haben also drastische Folgen für zeitliche Information einer Ereignissequenz, insbesondere für Arbeitslosigkeiten und Nichterwerbstätigkeiten.

#### *Episodenfrequenz und Dauerverteilung*

In Tabelle 5.7 werden die Interdependenzen näher beleuchtet, indem Veränderungen der Episodenfrequenz und der Dauerverteilungen für Episoden eines Typs einander gegenübergestellt werden. Eine Zu- oder Abnahme von Erwerbsepisoden hat in über 70 Prozent der Fälle *keine* Änderung der Anzahl der Monate in Erwerbstätigkeit zur Folge. Bei gleichbleibender Anzahl von Erwerbsepisoden verändert sich die Anzahl der Monate in Erwerbstätigkeit in etwa 12 Prozent der Fälle. Bei einer Ab- bzw. Zunahme von Arbeitslosigkeitsepisoden hingegen ergibt sich in weit über 80 Prozent der Fälle auch eine Ab- bzw. Zunahme der Monate in Arbeitslosigkeit, und bei gleichbleibender Anzahl von Arbeitslosigkeiten verändert sich die Dauer in Arbeitslosigkeit nur in etwa 4,5 Prozent der Fälle. Für Nichterwerbstätigkeiten stellt sich die Sachlage ähnlich dar: Wenn die Anzahl der Nichterwerbstätigkeiten abweichend angegeben wird, ändert sich die Dauer in Nichterwerbstätigkeit in 72 Prozent der Fälle, und bei einer gleichbleibenden Anzahl von Nichterwerbstätigkeiten ändert sie sich in weniger als 3 Prozent der Fälle.

Dies zeigt, dass die Dauer in Erwerbstätigkeit vergleichsweise unabhängig von der Anzahl der rekonstruierten Erwerbstätigkeiten ist, während die Stabilität der Anzahl von Monaten in Arbeitslosigkeit und Nichterwerbstätigkeit in hohem Maße von der Stabilität der Episodenrekonstruktion abhängt. Dies ist darauf zurückzuführen, dass eine Änderung der Anzahl der Arbeitslosigkeiten und Nichterwerbstätigkeiten meistens bedeutet, dass statt keiner eine oder statt einer keine Episode angegeben wird. Die Anzahl der Erwerbsepisoden hingegen kann sich auch durch Fusion und Defusion verändern, wodurch die Dauerinformation erhalten bleibt.

Tabelle 5.7: Interdependenz von Fehlern bei Episodenstruktur und Dauerverteilung

Veränderung der Episodenstruktur	Veränderung in der Dauerverteilung				N
	=		-+		
	absolut	in %	absolut	in %	
<b>Anzahl Erwerbstätigkeiten</b>					
	Monate in Erwerbstätigkeit				
=	638	87,9	88	12,1	726
-+	239	73,9	84	36,2	323
N	877		172		1.049
<b>Anzahl Arbeitslosigkeiten</b>					
	Monate in Arbeitslosigkeit				
=	913	95,7	41	4,3	954
-+	11	11,8	84	89,2	95
N	924		125		1.049
<b>Anzahl Nichterwerbstätigkeiten</b>					
	Monate in Nichterwerbstätigkeit				
=	906	97,6	22	2,4	928
-+	14	28,0	107	72,0	121
	920		129		1.049

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

*Kompensatorische Veränderungen in der Dauerverteilung*

In Tabelle 5.8 ist dargestellt, wie Zu- und Abnahme von Monaten in Episoden des einen durch Monate in Episoden des anderen Typs kompensiert wird. Eine *Abnahme* der Monate in *Erwerbstätigkeit* wird in 66 Prozent der Fälle von einer Zunahme der Monate in Arbeitslosigkeit und in über 40 Prozent von einer Zunahme der Monate in Nichterwerbstätigkeit kompensiert – und das heißt fast immer, dass eine ganze Arbeitslosigkeits- bzw. Nichterwerbstätigkeitsepisode hinzugefügt und die Erwerbstätigkeiten entsprechend verringert oder gestaucht wurden. Eine *Zunahme* der Monate in Erwerbstätigkeit wird in 35,9 Prozent der Fälle von einer Abnahme der Monate in Arbeitslosigkeit und in 59,0 Prozent von einer Abnahme der Monate in Nichterwerbstätigkeit begleitet; das heißt fast immer, dass eine ganze Arbeitslosigkeits- bzw. Nichterwerbstätigkeitsepisode unterschlagen und Erwerbstätigkeiten eingefügt oder gestreckt wurden. Wo Monate in Erwerbstätigkeit abnehmen, geschieht dies also häufiger, um Platz für Arbeitslosigkeit zu schaffen, und wo sie zunehmen, werden dadurch eher reduzierte oder verschwundene Nichterwerbstätigkeiten ausgeglichen.

Tabelle 5.8: Interdependenz von Fehlern bei der Dauerverteilung

		Veränderung in der Dauerverteilung											
		Monate in ET				Monate in ALO				Monate in NET			
		–	=	+	<i>n</i>	–	=	+	<i>n</i>	–	=	+	<i>n</i>
Monate in ET	–				5	27	62	94		16	39	39	94
					5,3%	28,7%	66,0%		17,0%	41,5%	41,5%		
	=				8	852	17	877		15	853	9	877
					0,9%	97,1%	1,9%		1,7%	97,3%	1,0%		
Monate in ALO	+				28	45	5	78		46	28	4	78
					35,9%	57,7%	6,4%		59,0%	35,9%	5,1%		
	<i>n</i>				41	924	84		77	920	52		
Monate in NET	–	5	8	28	41					1	24	16	41
		12,2%	19,5%	68,3%						2,4%	58,5%	39,0%	
	=	27	852	45	924					42	855	27	924
		2,9%	92,2%	4,9%						4,5%	92,5%	2,9%	
Monate in NET	+	62	17	5	84					34	41	9	84
		73,8%	20,2%	6,0%						40,5%	48,8%	10,7%	
	<i>n</i>	94	877	78						77	920	52	
Monate in NET	–	16	15	46	77	1	42	34	77				
		20,8%	19,5%	59,7%		1,3%	54,5%	44,2%					
	=	39	853	28	920	24	855	41	920				
		4,2%	92,7%	3,0%		2,6%	92,9%	4,5%					
Monate in NET	+	39	9	4	52	16	27	9	52				
		75,0%	17,3%	7,7%		30,8%	51,9%	17,3%					
	<i>n</i>	94	877	78		41	924	84					

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

– Weniger Monate in der zweiten Befragung.

= Identische Anzahl von Monaten in beiden Befragungen (ein Monat Toleranz).

+ Mehr Monate in der zweiten Befragung.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

Eine *Abnahme* der Monate in *Arbeitslosigkeit* (also das Unterschlagen einer Episode) wird in 68,3 Prozent der Fälle durch eine Zunahme der Monate in Erwerbstätigkeit (also das Strecken der Standardaktivität) kompensiert und in 39 Prozent der Fälle durch eine Zunahme der Monate in Nichterwerbstätigkeit (also durch Umdefinition). Auch die Zunahme der Monate in Arbeitslosigkeit (also das Hinzufügen einer Episode) wird häufiger durch die Abnahme von Erwerbstätigkeit ausgeglichen.

Was *Nichterwerbstätigkeit* betrifft, so wird ihre *Abnahme* (also das Unterschlagen einer Episode) häufiger, nämlich in etwa 60 Prozent der Fälle, durch das Strecken der Erwerbstätigkeit und in 40 Prozent der Fälle von einer Zunahme von Arbeitslosigkeit ausgeglichen, also vermutlich durch Umdefinieren. Dasselbe gilt für die Zunahme von Nichterwerbstätigkeit, die in 75 Prozent aller Fälle durch eine Zunahme in Erwerbstätigkeit und in 30 Prozent durch eine Zunahme von Arbeitslosigkeit kompensiert wird.

Wer also weniger oder mehr Arbeitslosigkeiten oder Nichterwerbstätigkeiten angibt, gleicht dies in der Sequenz am häufigsten durch das Strecken oder Stauchen von Erwerbstätigkeit. In geringerem Umfang kommt es auch dazu, dass das eine in das andere umgedeutet oder auf Kosten des jeweils anderen reduziert wird. Besonders die Zunahme der Monate in Arbeitslosigkeit oder Nichterwerbstätigkeit geht vor allem zu Lasten von Erwerbstätigkeit.

### 5.3.6 Folgen für den ersten Übergang und sein Timing

Nun werde ich untersuchen, wie sich das Timing des ersten Übergangs durch Gedächtnisfehler verändert. Ich verwende hier nur die 949 Personen (90,5%), die bei Schritt 1 als erste Episode im Zeitfenster übereinstimmend eine Erwerbstätigkeit angegeben haben. Da bei Schritt 2 nicht jedes Mal ein Übergang im Zeitfenster rekonstruiert wurde, gibt es vier Möglichkeiten (siehe Abb. 5.6): Sie endet beide Male nicht im Zeitfenster, sie endet nur in einer der Befragungen im Zeitfenster oder sie endet beide Male im Zeitfenster. Vollständig können Unterschiede im Timing nur für die 366 Übergänge beschrieben werden, die in beiden Befragungen im Zeitfenster liegen (Fall d). Wenn in nur einer der beiden Befragungen ein Übergang im Zeitfenster angegeben wird, ist zumindest die Richtung des Unterschieds bekannt: Im Fall b findet der Übergang offensichtlich in der zweiten Befragung später statt, in Fall c hingegen findet er in der ersten Befragung später statt.

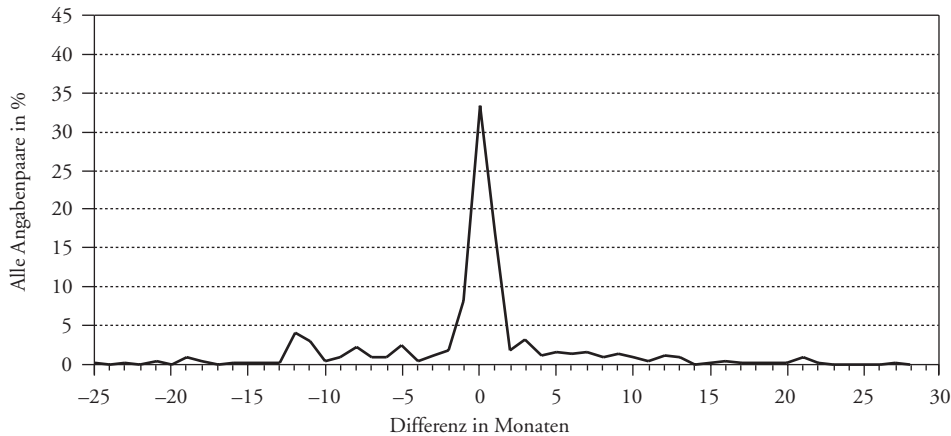
Abbildung 5.6: Timing des ersten Übergangs aus Erwerbstätigkeit im Zeitfenster

		<i>N</i>	%
Fall a	In beiden Befragungen kein Übergang im Zeitfenster	369	38,9
Fall b	Nur erste Befragung Übergang im Zeitfenster	149	15,7
Fall c	Nur zweite Befragung Übergang im Zeitfenster	65	6,8
Fall d	In beiden Befragungen Übergang im Zeitfenster	366	38,9
Insgesamt		1.049	100,0

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.



Abbildung 5.7: Datierung des ersten Übergangs aus Erwerbstätigkeit – Verteilung der Abweichungen



Positive Werte: Datum in der zweiten Befragung später; negative Werte: Datum in der zweiten Befragung früher.  
 $n = 366$  datierte Übergänge.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

Abbildung 5.7 zeigt für die 366 Übergänge, für die in beiden Befragungen ein Datum innerhalb des Zeitfensters vorliegt, die Verteilung der Abweichungen in Monaten. Positive Abweichungen stehen für einen späteren Übergang und negative Abweichungen für einen früheren Übergang in der zweiten Befragung. Irrtumsrisiko und Größe der Abweichungen sind, wie für inferenziell rekonstruierte Daten zu erwarten, beträchtlich: 41,5 Prozent der Datumsangaben liegen mehr als einen Monat auseinander, etwa 25 Prozent mehr als ein halbes Jahr und 7 Prozent mehr als 12 Monate. Die charakteristische Häufung bei etwa einem Jahr, die auf einen Skaleneffekt beim Datieren hinweist, ist deutlich sichtbar – aber nur, wenn in der zweiten Befragung ein früheres Datum angegeben wird als in der ersten. Dies liegt vermutlich daran, dass das Grenzsteinereignis des Mauerfalls die Zuordnung zum korrekten Jahr zumindest für das erste Jahr des Zeitfensters erleichtert; Grenzsteine sind umso wirksamer, je näher ihnen das zu datierende Ereignis ist (Auriat, 1996). Abgesehen davon ist die Richtung des Fehlers innerhalb des Zeitfensters symmetrisch.

Um die Abweichungen im Timing, die auf die Rekonstruktion unterschiedlicher Übergänge zurückzuführen sind, von denen zu unterscheiden, die tatsächlich beim Datieren desselben Übergangs entstehen, betrachte ich nun die konsistent und inkonsistent beschriebenen Übergänge getrennt. Für inkonsistent beschriebene erste Übergänge erwarte ich, dass die Abweichungen häufiger auftreten, größer sind und sich das Timing häufiger nach hinten verschiebt<sup>7</sup>.

<sup>7</sup> Zwar ist nicht davon auszugehen, dass jeder Übergang, der konsistent beschrieben wurde, sich auch wirklich auf dasselbe Ereignis bezieht (siehe z.B. Einzelfall 11c). Umgekehrt wird sich auch nicht jeder abweichend beschriebene Übergang auf ein gänzlich anderes Ereignis beziehen (siehe z.B. Einzelfall 8a, 9a und 9b). Allerdings ist anzunehmen, dass der Anteil äquivalenter Ereignisse in der Gruppe derer mit konsistent beschriebenen Übergängen deutlich höher ist.

Tabelle 5.9: Erster Übergang aus Erwerbstätigkeit – Veränderungen zwischen erster und zweiter Befragung

Erste Befragung	Kein Übergang im Zeitfenster		ET-ET inner- betrieblich		ET-ET zwischen- betrieblich		ET-ALO		ET-NET		Gesamt	
	absolut	in%	absolut	in%	absolut	in%	absolut	in%	absolut	in%	absolut	in%
	Kein Übergang im Zeitfenster	369	85,0	32	7,4	18	4,1	10	2,3	5	1,1	434
ET-ET inner- betrieblich	61	49,6	33	26,8	17	13,8	7	5,7	5	4,1	123	12,9
ET-ET zwischen- betrieblich	69	26,8	31	12,1	126	49,0	16	6,2	15	5,8	257	27,1
ET-ALO	9	10,8	4	4,8	3	3,6	63	75,9	4	4,8	83	8,7
ET-NET	10	19,2	4	7,7	2	3,8	21	40,4	15	28,8	52	5,5
Gesamt	518	54,6	104	10,9	166	17,5	117	12,3	44	4,6	949	

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

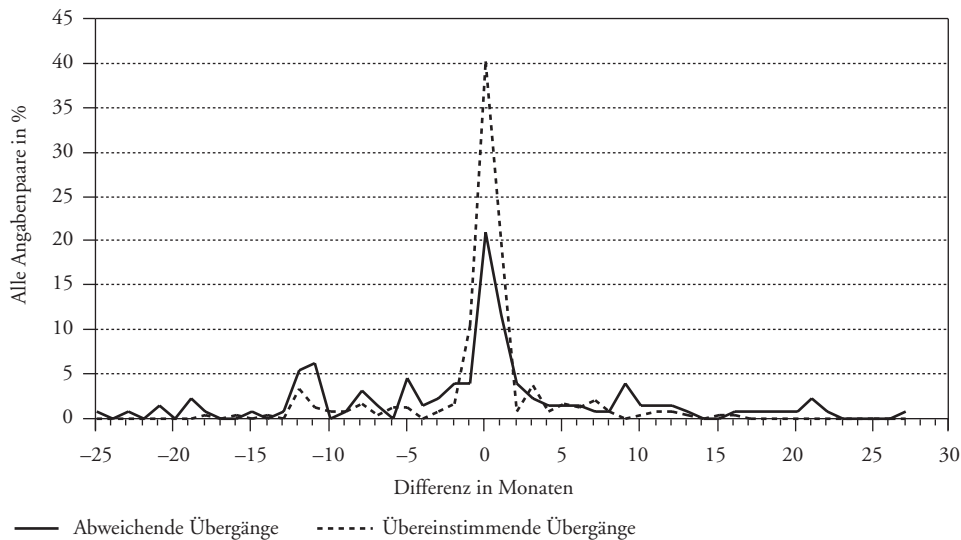
Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

Tabelle 5.9 zeigt zunächst für alle 949 Lebensläufe, welche ersten Übergänge abweichend rekonstruiert werden. Am stabilsten sind die Angaben, wenn im gesamten Zeitfenster kein Übergang stattgefunden hat. Ansonsten werden wie erwartet, Übergänge in Arbeitslosigkeit am häufigsten konsistent angegeben, gefolgt von zwischenbetrieblichen Wechseln, während innerbetriebliche Wechsel und Wechsel in Nichterwerbstätigkeit beide in etwa drei Viertel der Fälle nicht konsistent angegeben werden. Dass in der ersten Befragung ein inner- und in der zweiten Befragung ein zwischenbetrieblicher Wechsel angegeben wird, kommt im selben Umfang vor wie die umgekehrte Abweichung.

In Abbildung 5.8 ist die Abweichung der Datumsangaben in Monaten getrennt für konsistent und inkonsistent rekonstruierte Übergänge angegeben. Erwartungsgemäß haben die konsistent benannten ersten Übergänge ein deutlich geringeres Risiko, fehldatiert zu werden, und große Abweichungen von mehr als einem Jahr kommen seltener vor. Eine Neigung, das Timing bei abweichend rekonstruierten Übergängen deutlich häufiger nach hinten zu verlegen, ist nicht zu sehen; die Abweichungen verteilen sich sowohl für konsistent als auch für inkonsistent benannte Übergänge weitestgehend symmetrisch.

Diese Symmetrie verschwindet allerdings bei Einbeziehung derjenigen Personen, die nur in einer der beiden Befragungen einen Übergang berichten (siehe Abb. 5.6). Fall b, in dem nur in der ersten Befragung ein Übergang angegeben und das Timing des ersten Übergangs nach hinten verlegt wird, tritt mehr als doppelt so häufig auf wie Fall c, in dem das Timing nach vorne verlegt wird. Ich vermute, dass hier meist nicht nur derselbe Übergang bei Schritt 3 versehentlich aus dem Zeitfenster hinaus- bzw. in das Zeitfenster hineindatiert wurde, sondern dass bei Schritt 2 der Übergang unterschlagen oder hinzugefügt wurde. Da Übergänge in der zweiten Befragung häufiger unterschlagen als neu eingefügt werden – bzw. Fusionierungen häufiger vorkommen als Defusionierungen –, verlängern sich Episoden häufiger.

Abbildung 5.8: Datierung des ersten Übergangs aus Erwerbstätigkeit im Zeitfenster für übereinstimmend und abweichend benannte Übergänge



Positive Werte: Datum in der zweiten Befragung später; negative Werte: Datum in der zweiten Befragung früher.  
 $n = 366$  datierte Übergänge.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

ger, als dass sie sich verkürzen, und das Ende einer Episode wird in der zweiten Befragung häufiger aus dem Zeitfenster hinaus- als hineinverlegt. Dies deutet also darauf hin, dass die Inkonsistenzen im Timing, die auf abweichende Rekonstruktion bei Schritt 2 des Datengenerierungsprozesses zurückgehen, ausfallen, während der Fehler durch das Datieren bei Schritt 3 symmetrisch ausfällt.

Bei der wiederholten Rekonstruktion der Erwerbssequenzen haben sich zahlreiche Veränderungen in der Episodenstruktur ergeben, die in alle Richtungen auftreten. Abgesehen von der Zunahme von Arbeitslosigkeit kommt es dadurch insgesamt zu einer Vereinfachung und Konventionalisierung der Sequenzen. Frauen und Männer sowie die befragten Kohorten weisen dabei einige Unterschiede in der Konsistenz auf; keine der Gruppen ist dabei jedoch bei allen Episodentypen und allen Fehleraspekten über- oder unterlegen. Die Abweichungen treten auch abhängig vom tatsächlichen Sachverhalt auf. Fehler beim Datieren haben vergleichsweise geringe Auswirkungen, aber durch Abweichungen bei Schritt 1 und 2 der Datengenerierung werden die Episodenstruktur und damit die Dauerverteilungen sowie das Timing des ersten Übergangs systematisch verändert. Im folgenden Abschnitt werde ich betrachten, wie sich diese Veränderungen auf die Schlussfolgerungen über die Nachwendeerwerbsverläufe niederschlagen können.

## 5.3.7 Folgen für die inhaltlichen Auswertungen

Um mögliche Folgen der beschriebenen Abweichungen auf die Ergebnisse von deskriptiven und explanatorischen Analysen abzuschätzen, werde ich erstens einige ausgewählte deskriptive Informationen miteinander vergleichen und zweitens ein logistisches Regressionsmodell einmal mit den Daten aus der ersten und einmal mit den Daten aus der zweiten Befragung berechnen.

Bei der Wahl der zu vergleichenden deskriptiven Informationsdimensionen orientiere ich mich an denen, die Mayer (2001) verwendet, um die soziale Mobilität Ostdeutscher im Transformationszeitraum zu beschreiben. Mayer konstatiert hier: „Bis in das erste Halbjahr 1996 verliessen etwa 70 % unserer Befragten die Arbeitsstätte, in der sie 1989 gearbeitet hatten.“ An anderer Stelle stellt er fest, dass „direkte Firmenwechsel ohne Zwischenarbeitslosigkeit in der Periode nach 1989 in der Tat höher waren als Freisetzungen in Arbeitslosigkeit“. Diese deskriptiven Aussagen werde ich nun getrennt mit den Daten der ersten und der zweiten Befragung für den Zeitraum von Dezember 1989 bis Dezember 1991 – knapp zwei Jahre nach dem Mauerfall – überprüfen. Ich werde Männer und Frauen sowie die drei Kohorten getrennt betrachten, um die Folgen der teilweise unterschiedlichen Erinnerungsgüte dieser Gruppen abschätzen zu können<sup>8</sup>.

Tabelle 5.10 zeigt für diejenigen, die im Dezember 1989 erwerbstätig waren (in der ersten Befragung 973, in der zweiten 979 Personen) den Anteil derjenigen, die im Dezember 1991 noch im selben Betrieb tätig waren, die also weder in einen anderen Betrieb, noch in Arbeitslosigkeit oder Nichterwerbstätigkeit übergetreten waren. Dieser Anteil schwankt in beiden Befragungen nur um ein bis zwei Prozentpunkte, teilweise sogar nur im Nachkommabereich. Auch die Geschlechter- und Kohortenunterschiede stellen sich in den beiden Befragungen sehr ähnlich dar.

Tabelle 5.10: Stabilität der Betriebszugehörigkeit in der ersten und zweiten Befragung

	Erste Befragung		Zweite Befragung	
	Noch im selben Betrieb		Noch im selben Betrieb	
	absolut	in %	absolut	in %
Alle	973	100,0	979	100,0
Männer	480	49,3	481	49,1
Frauen	493	50,7	498	50,9
Kohorte um				
1940	340	35,0	338	34,5
1950	320	32,9	314	32,0
1960	313	32,0	327	33,4

Anteil der Befragten, die im Dezember 1991 noch im selben Betrieb tätig waren wie im Dezember 1989.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen

<sup>8</sup> Die dabei ermittelten Zahlen können mit denen der Studie von Mayer nicht verglichen werden, da dort (a) die edierten Daten verwendet werden (mehr dazu in Abschnitt 5.3.8) und (b) zumindest bei der ersten Frage ein Zeitraum von 1989 bis 1996 betrachtet wird, während das Zeitfenster hier nur bis Mitte 1992 reicht.

Tabelle 5.11 zeigt, wie häufig bis Dezember 1991 zwischenbetriebliche Wechsel, Wechsel in Arbeitslosigkeit oder in Nichterwerbstätigkeit auftraten. Sowohl in der ersten als auch in der zweiten Befragung sind die Übergänge in einen anderen Betrieb in diesem Zeitraum häufiger als Übergänge in Arbeitslosigkeit. Der Abstand fällt in der zweiten Befragung allerdings deutlich schwächer aus, und für Frauen sind in der zweiten Befragung Wechsel in Arbeitslosigkeit sogar häufiger als Betriebswechsel. Die Kohortenunterschiede stellen sich in den beiden Befragungen größtenteils ähnlich dar; in der zweiten Befragung fallen die Unterschiede aber für alle Kohorten schwächer aus und für die Kohorte der um 1950 Geborenen sogar so schwach, dass in der zweiten Befragung Übergänge in Arbeitslosigkeit und zwischenbetriebliche Wechsel etwa gleich häufig auftreten.

Deskriptive Information über Zustands- und Übergangsverteilungen können also durch Erinnerungsfehler beeinflusst werden, dies ist aber nicht immer und nicht für alle Befragten-Gruppen der Fall.

Um die Einflüsse der abweichenden Rekonstruktionen auf Beziehungen von Variablen untereinander zu beleuchten, werde ich ein einfaches logistisches Regressionsmodell aufstellen, das das Risiko, aus der für Dezember 1989 angegebenen Erwerbstätigkeit in eine Arbeitslosigkeit oder Nichterwerbstätigkeit zu wechseln, aus Geschlecht, Kohortenzugehörigkeit, Schul- und Ausbildungsabschluss vorhersagt. Außerdem nehme ich den Zeitpunkt, zu dem der Übergang auftritt, als kontinuierliche Erklärungsvariable mit auf: Das Risiko eines Übergangs in Arbeitslosigkeit oder Nichterwerbstätigkeit sollte umso höher liegen, je später im Zeitfenster der Übergang stattfindet, da Massenentlassungen erst ab Mitte 1990 stattfanden (Goedicke, 2000)<sup>9</sup>. Ich berücksichtige alle Personen, die in der jeweiligen Befragung im Dezember 1989 erwerbstätig waren und bis einschließlich Dezember 1991 einen Übergang berichteten.

Tabelle 5.11: Übergänge aus Erwerbstätigkeit bis Dezember 1991 in der ersten und zweiten Befragung (in %)

	Erste Befragung			Zweite Befragung				
	<i>N</i>	Übergänge in andere ET (mit Betriebswechsel)	Übergänge in ALO	Übergänge in NET	<i>N</i>	Übergänge in andere ET (mit Betriebswechsel)	Übergänge in ALO	Übergänge in NET
Alle	424	63,4	22,4	14,2	360	49,7	35,4	14,7
Männer	219	74,4	19,6	5,9	189	60,3	26,5	13,2
Frauen	205	51,7	25,4	22,9	171	38,0	45,2	16,4
Kohorte um								
1940	126	58,7	24,3	16,7	102	43,1	42,6	14,7
1950	142	64,8	22,5	12,7	115	55,7	35,7	8,7
1960	153	66,0	20,5	13,5	133	49,6	31,6	18,8

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

<sup>9</sup> Dieses Modell ist sicherlich ungeeignet, die Erwerbskarrieren der Befragten im Nachwendezeitraum angemessen zu erklären. Allerdings kann es Aufschluss darüber geben, ob und wie die Modellkoeffizienten sich zwischen den beiden Befragungen unterscheiden können.

In Tabelle 5.12 sind in den beiden linken Spalten die Modellkoeffizienten  $\exp(b)$  dargestellt, wie sie sich mit den Daten der ersten und mit den Daten der zweiten Befragung ergeben.  $\exp(b)$  wird als die relative Zunahme des Risikos, in Arbeitslosigkeit oder Nichterwerbstätigkeit zu wechseln, gegenüber einer Vergleichsgruppe interpretiert. Die Koeffizienten ändern sich geringfügig in ihrer Höhe; in der zweiten Befragung beispielsweise erhöht ein Fachschulabschluss das Risiko, in eine Arbeitslosigkeit oder Nichterwerbstätigkeit zu wechseln, stärker als in der ersten Befragung. Ihre Wirkrichtung bleibt aber immer gleich, und die Koeffizienten, die in einem Modell signifikant sind, sind es auch im anderen.

In den beiden rechten Spalten von Tabelle 5.12 rechne ich dasselbe Modell noch einmal und berücksichtige dabei auch diejenigen, die bis Dezember 1991 keinen Übergang berichteten. Für diese setze ich das Ende des Zeitfensters als Datum des Übergangs ein. Hier fallen die Koeffizienten diskrepanter aus. Der Koeffizient für Hochschulabsolvanz ist in der zweiten Befragung nicht mehr signifikant. Die Koeffizienten für Geschlecht und Fachschulabschluss sind beide Male signifikant, ändern aber ihre Höhe deutlich. Der Koeffizient für die verstrichene Zeit ändert sogar die Richtung: In der ersten Befragung sinkt die Chance, in Beschäftigung zu bleiben, mit jedem verstrichenen Monat, in der zweiten hingegen steigt sie.

Tabelle 5.12: Determinanten des Risikos für Erwerbsverlust in der ersten und zweiten Befragung

	Erste Befragung	Zweite Befragung	Erste Befragung	Zweite Befragung
	(nur Übergänge im Zeitfenster)	(nur Übergänge im Zeitfenster)	(auch zensierte)	(auch zensierte)
	Exp(b)	Exp(b)	Exp(b)	Exp(b)
(Referenzkategorie: Männlich)				
Weiblich	.328***	.378***	.407***	.510***
(Referenzkategorie: 1940er Kohorte)				
1950er Kohorte	(n.s.)	(n.s.)	(n.s.)	(n.s.)
1960er Kohorte	(n.s.)	(n.s.)	(n.s.)	(n.s.)
(Referenzkategorie: Volksschule)				
POS	(n.s.)	(n.s.)	(n.s.)	(n.s.)
EOS und Abitur	(n.s.)	(n.s.)	(n.s.)	(n.s.)
(Referenzkategorie: Ungelernte und Facharbeiter)				
Meister	(n.s.)	(n.s.)	(n.s.)	(n.s.)
Fachschulabschluss	1,947*	2,719***	1,915*	2,208**
Hochschulabschluss	3,920*	3,702***	3,260*	(n.s.)
Monate seit Dezember 1989 (kontinuierliche Variable)	(n.s.)	(n.s.)	.407*	1,156***
<i>N</i>	404	335	979	921
Konstante	30,01 (n.s.)	24,10 (n.s.)	-120,978***	-157,563***
-1 Log-Likelihood	479,93	423,801	670,94	670,73
Nagelkerke	0,142	0,151	0,211	0,302

Logistische Regression für das Risiko, im Zeitfenster in Arbeitslosigkeit oder Nichterwerbstätigkeit zu wechseln.

Signifikanzniveau: \*  $p < 0,1$ , \*\*  $p < 0,05$ , \*\*\*  $p < 0,01$ ; n.s. = nicht signifikant.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.

Modellparameter können also in ihrer Größe, Richtung und Signifikanz von den Unterschieden in den Datensätzen beeinflusst werden. Die anfälligste Größe, die sogar ihre Wirkrichtung änderte, war dabei die Variable, in der der Zeitverlauf kodiert war, nämlich die Monate seit Dezember 1989. Diese Beeinflussung fällt angesichts der verbreiteten, großen und systematischen Abweichungen der rekonstruierten Sequenzen allerdings erstaunlich moderat aus.

### 5.3.8 Verbesserung der Übereinstimmung durch Datenedition

Um die Erfolge der Datenedition bei der Reduktion von Gedächtnisfehlern abzuschätzen, werde ich nun einige der bisher angestellten Vergleiche mit den edierten Daten der zweiten Befragung wiederholen. In Tabelle 5.13 sind die Übereinstimmungsraten auf Individualebene für die Zusammensetzung der Erwerbssequenzen getrennt für den edierten und unedierten Datensatz aufgelistet, in Tabelle 5.14 für die ersten Übergänge, und in Tabelle 5.15 werden die Anteile äquivalent rekonstruierter Sequenzen verglichen. Nahezu überall verbessert sich die Übereinstimmung durch die Edition und dies teilweise erheblich: Für zwischenbetriebliche Wechsel beispielsweise steigt die Übereinstimmung von 49 auf über 79 Prozent (Tab. 5.15) und für Erwerbsverläufe mit Episoden aller drei Typen von 38,5 auf 73,3 Prozent (Tab. 5.14). Insbesondere steigt der Anteil äquivalent rekonstruierter Sequenzen für diejenigen mit mehr als nur einer Episode, das heißt, dass die Abhängigkeit der Abweichungen vom tatsächlichen Sachverhalt in der Edition reduziert wird.

Allerdings gibt es einzelne Abweichungen, die sich durch die Edition nicht oder nur wenig verbessert haben. Beispielsweise liegt der Anteil derjenigen, die in der ersten Befragung einen zwischenbetrieblichen Wechsel als ihren ersten Übergang angeben, in der zweiten aber einen innerbetrieblichen, im edierten wie im unedierten Datensatz bei etwa 12 Prozent. Auch der Anteil der Befragten, die statt einer Sequenz aus Erwerbstätigkeiten und Nichterwerbstätigkeiten eine aus Erwerbstätigkeiten und Arbeitslosigkeit angeben, liegt im edierten wie im unedierten Datensatz bei 19 Prozent. Und der Anteil der Befragten, die statt eines innerbetrieblichen Wechsels als ersten Übergang einen zwischenbetrieblichen angeben, ist durch die Edition sogar von 14 auf 20,3 Prozent gestiegen. Am wenigsten verbesserten sich also die Abweichungen, die in geringerem Maße zu Widersprüchen und Inkonsistenzen mit anderen Bereichen führen, sodass sie in der Edition seltener überhaupt bemerkt werden.

## 5.4 Diskussion

### *Rekonstruktionsveränderungen durch die verlängerte retrospektive Distanz und ihre Folgen: Zusammenfassung*

In diesem Kapitel habe ich gezeigt, dass der dreistufige Rekonstruktionsprozess, den die Befragten in der LV vollführen, in der zweiten Befragung häufig anders abläuft als in der ersten. Bei der wiederholten Rekonstruktion der Erwerbssequenzen haben sich zahlreiche Veränderungen in der Episodenstruktur ergeben. Abgesehen von der Zunahme von Arbeitslosigkeit laufen diese auf meist eine Vereinfachung und Konventionalisierung der Sequenzen hinaus, es kommt jedoch auch in beträchtlichem Umfang dazu, dass Sequenzen in der län-

Tabelle 5.13: Zusammensetzung und Komplexität der Erwerbsequenzen – Übereinstimmung der unedierten und edierten Daten

	Erste Befragung				Zweite Befragung				Episoden aller drei Typen				Gesamt								
	Nur Episoden eines Typs		ET und ALO		ET und NET		ALO und NET		unedierte		ediert		unedierte		ediert						
	abs.	in %	abs.	in %	abs.	in %	abs.	in %	abs.	in %	abs.	in %	abs.	in %	abs.	in %					
Nur Episoden eines Typs	765	92,8	797	96,8	30	3,7	14	1,7	27	3,3	10	1,2	1	1	1	1	823	78,5			
ET und ALO	18	17,8	3	3,0	67	66,3	83	82,2	5	4,9	4	4,0	3	3,0	8	7,9	11	10,9	101	9,6	
ET und NET	38	39,1	26	26,8	19	19,6	19	19,6	30	30,9	46	44,3	3	3,1	7	7,2	9	9,3	97	9,2	
ALO und NET													1	1	1	1	1	1	2		
Episoden aller drei Typen	1				9	32,1	7	32,1	5	7,5			1		10	38,5	19	73,3	26	2,5	
Gesamt	821	78,3	826	78,7	125	11,9	123	11,7	67	6,4	57	5,4	9	0,9	2	0,2	27	2,6	41	3,9	1.049

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin; eigene Berechnungen.



Tabelle 5.14: Erster Übergang aus Erwerbstätigkeit – Übereinstimmung der unedierten und edierten Daten

Erste Befragung	Kein Übergang im Zeitfenster				Zweite Befragung				N														
	ET-ET innerbetrieblich		ET-ET zwischenbetrieblich		ET-ALO		ET-NET																
	unediert abs.	ediert in %	unediert abs.	ediert in %	unediert abs.	ediert in %	unediert abs.	ediert in %															
Kein Übergang im Zeitfenster	391	85,9	408	89,7	32	7,0	20	4,4	16	3,5	16	3,5	10	2,2	6	1,8	5	1,1	5	1,1	455	455	
ET-ET inner- betrieblich	61	50,4	46	37,4	33	27,3	48	39,0	17	14,0	25	20,3	7	5,8	4	3,3	3	2,5			121	123	
ET-ET zwi- schenbetrieblich	69	26,8	29	11,3	31	12,1	30	11,5	126	49,0	190	72,8	16	6,2	8	3,1	15	5,8	4		1,5	257	261
ET-ALO	9	10,8	1		4	4,8	2		3	3,6	2		63	75,9	83	90,2	4	4,8	6	6,5	83	92	
ET-NET	10	19,6	7	12,7	3	5,9	2	3,6	2	3,9	2	3,6	21	41,2	25	45,5	15	29,4	19	34,5	51	55	
N	540		493		103		101		164		234		117		126		42		34		967	986	

ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit; NET = Nichterwerbstätigkeit.

Berücksichtigt sind alle Befragten, die im jeweiligen Datensatz als erste Episode eine Erwerbstätigkeit angeben.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung; eigene Berechnungen.

Tabelle 5.15: Äquivalente Sequenzen in den unedierten und edierten Daten

Äquivalente Sequenzen	Alle		1		2		3		4 bis 7										
	unediert abs.	ediert in %	unediert abs.	ediert in %	unediert abs.	ediert in %	unediert abs.	ediert in %	unediert abs.	ediert in %									
	N		N		N		N		N										
589	56,1	724	69,0	393	84,3	409	87,8	155	39,0	235	59,2	35	25,4	63	54,3	6	12,2	17	34,7
N	1.049		1.049		466		466		397		397		138		138		49		49

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung; eigene Berechnungen.

geren Rückschau wechselhafter rekonstruiert und Übergänge und Episoden hinzugefügt werden.

Die resultierenden Veränderungen lassen sich als Unterschlagen, Hinzufügen, Strecken, Stauchen, Umdefinieren, Fusionieren und Defusionieren von Episoden bzw. als Unterschlagen, Hinzufügen, Umdefinieren und abweichend Datieren von Übergängen beschreiben. Die Abweichungen treten auch abhängig vom tatsächlichen Sachverhalt auf. Besonders betroffen sind wie erwartet Arbeitslosigkeit und Nichterwerbstätigkeit, von den Übergängen die innerbetrieblichen Wechsel und von den Personen diejenigen mit komplexen, wechselhaften und episodischen Erwerbsverläufen. Frauen und Männer sowie die befragten Kohorten weisen dabei einige Unterschiede in der Konsistenz auf; keine der Gruppen ist dabei jedoch bei allen Episodentypen und allen Fehleraspekten unterlegen. Als Folge der Veränderungen verbringen die Befragten tendenziell mehr Zeit in Arbeitslosigkeit und weniger in Nichterwerbstätigkeit, während die Zeit in Erwerbstätigkeit sich weniger stark verändert. Das Timing des ersten Übergangs wird tendenziell nach hinten verlegt.

Inhaltliche Schlussfolgerungen können anders ausfallen und Modellparameter in ihrer Signifikanz, Höhe und Wirkrichtung beeinflusst werden. Wie in Pierrets (2001) Modellvergleich waren dabei die Variablen am labilsten, in denen Information über zeitliche An- und Einordnung von Ereignissen enthalten war. Diese Veränderungen fallen angesichts der verbreiteten und systematischen Abweichungen allerdings erstaunlich moderat aus. In der Edition wurden viele der gedächtnisverursachten Inkonsistenzen jedoch gefunden und korrigiert, wodurch sich die Übereinstimmung in vielen Punkten erheblich verbesserte. Relativ wenig konnten jedoch Irrtümer bei der Rekonstruktion innerbetrieblicher Wechsel und das Umdefinieren von Arbeitslosigkeitsphasen in Nichterwerbstätigkeiten (und umgekehrt) korrigiert werden.

#### *„Ostspezifische“ Abweichungen*

Der tief greifende und umfassende gesellschaftliche Wandel im Nachwendzeitraum schafft besondere Bedingungen für die Rekonstruktion von Erwerbsverläufen. Schon von der Menge her treten vergleichsweise viele rasch aufeinander folgende Veränderungen auf. Auch treten Zeiträume, in denen die Aktivitäten potenziell mehrdeutig sind, wesentlich häufiger auf, etwa Kurzarbeit oder Arbeitsbeschaffungsmaßnahmen. Durch die Instabilität der betrieblichen Struktur wird zusätzlich die Unterscheidung von zwischen- und innerbetrieblichen Wechseln uneindeutig. Die Zusammenfügung der Erwerbsverläufe zu einer narrativen und zielorientierten Sequenz gestaltet sich außerdem schwierig, da die Erwerbsverläufe im fraglichen Zeitraum vielfach durch die unkontrollierbaren und unvorhersehbaren äußeren Umwälzungen bestimmt waren. Außerdem geht die längere retrospektive Distanz in der zweiten Befragung durch den radikalen Wandel mit grundlegend veränderten Erinnerungskontexten einher. Als Resultat kam es, im Widerspruch zu Befunden aus anderen Studien, zu einer Zunahme von Arbeitslosigkeitsepisoden und zu einem Mangel an Geschlechterunterschieden bei der Erinnerung an Arbeitslosigkeit und Erwerbstätigkeit.

Zwar stellt sowohl in Ost- als auch in Westdeutschland Arbeitslosigkeit ein stigmatisiertes Ereignis dar, das sich nur negativ als Abwesenheit von Erwerbsarbeit definiert. Mit der kollektiven Erfahrung der (schuldlosen) Massenarbeitslosigkeit jedoch, die sich im Osten Deutschlands durch die lang anhaltende Beschäftigungskrise ergibt, wird dieses Stigma reduziert.

Wenn Arbeitslosigkeit ein Bestandteil der „Normalbiografie“ geworden ist, muss ihr Auftreten weniger biografisch begründet werden als in stabilen Beschäftigungssystemen mit vergleichsweise niedrigen Arbeitslosenquoten und mehr Kontrolle über Arbeitsplatzwechsel (Mutz, 1996). Insofern kann die gestiegene Bereitschaft, Zeiträume als Arbeitslosigkeit zu beschreiben, als Anpassung der Rekonstruktion an die veränderte individuelle und normalbiografische Lebenslogik begriffen werden.

Das Ausbleiben von Geschlechtereffekten bei der Erinnerung an Erwerbstätigkeit und Arbeitslosigkeit spiegelt die Sozialisation der Befragten wider. Da in der DDR weibliche Erwerbsarbeit als Selbstverständlichkeit oder Errungenschaft angesehen wurde, ergeben sich unterschiedliche Wertigkeiten und Deutungsmuster von Erwerbsarbeit ostdeutscher Frauen. Diese weisen auch noch Jahre nach der Wende eine ungebrochene Erwerbsorientierung auf und haben sich nicht als Reaktion auf die schwierige Arbeitsmarktsituation in die Familienarbeit zurückgezogen (Berger, 1995; Harre & Schmidt, 1995; Nickel, 1994). Die überlegene Erinnerung an die Nichterwerbstätigkeiten spiegelt wider, dass trotz verbreiteter weiblicher Vollzeit-erwerbstätigkeit auch in der DDR Familien- und Hausarbeit in die Zuständigkeit der Frauen fiel. Insofern spielt auch diese Sphäre bei der Selbstwahrnehmung und -konstruktion eine stärkere Rolle.

Der historische Wandel hat für alle Altersgruppen zeitgleich unsichere und wechselhafte Erwerbsbedingungen geschaffen; die Erwartung, dass sich dies im Fehlen von Kohortenunterschieden niederschlagen würde, konnte allerdings nicht bestätigt werden. Die jüngeren Teilnehmer waren inkonsistenter bei Erwerbstätigkeiten und konsistenter bei Arbeitslosigkeiten und Nichterwerbstätigkeiten. Möglicherweise haben die älteren Teilnehmer gerade wegen ihrer vergleichsweise wenig wechselhaften Erwerbsverläufe und längeren Sozialisation im Beschäftigungssystem der DDR besondere Schwierigkeiten, Arbeitslosigkeit und Nichterwerbstätigkeit in die Logik ihrer Erwerbsgeschichte einzuordnen, während dies den jüngeren Kohorten besser gelingt. Hier wären spezifischere Hypothesen über die differenzielle Wirkung der Veränderungen auf die drei Geburtsjahrgänge nötig, um die Muster in den Unterschieden schlüssig zu erklären. Auch mögliche Interaktionseffekte von Kohorte und Geschlecht sollten dabei ins Auge gefasst werden (Elias, 1991; Paull, 2002).

Die Besonderheiten der Ergebnisse aus Ostdeutschland sind mehr als ein historisches Kuriosum. Erstens zeigen sie, dass in standardisierten Lebensverlaufsbefragungen trotz ihres geringen Interpretationsspielraums der institutionelle, historische und individuelle Kontext die Rekonstruktion autobiografischer Information mitbestimmt. Diese Kontexte gestalten einerseits die individuellen Erfahrungen, die Individuen machen und beeinflussen andererseits individuelle und normative Deutungslogiken. Deshalb können dieselben Mechanismen der Informationsverarbeitung in verschiedenen Kontexten zu unterschiedlichen Datenfehlern führen.

Zweitens gelten diese Befunde generell für im Wandel befindliche und krisengeprägte Arbeitsmärkte und Beschäftigungssysteme. In dem Maße, in dem die lebenslange Vollzeit-erwerbstätigkeit auch in Westdeutschland und anderen modernen Industriegesellschaften weniger typisch wird (Berger & Sopp, 1992; Kohli, 1985; Mayer, 1998), verändern sich auch die normativen schematischen Erwartungen für Lebens- und Erwerbsverläufe. Und auch in weitgehend stabilen Gesellschaften kann dies für bestimmte Gruppen der Fall sein, etwa für die weniger gut Ausgebildeten oder die Bewohner strukturschwacher Regionen. Dort wo potenziell zweideutige Zustände wie Kurzarbeit, Arbeitsbeschaffungsmaßnahmen und auch Arbeitslosigkeit verbreitet sind, muss deren angemessene Erfassung also stärker bedacht werden.

*Datengenerierung und Datenedition – Wo entstehen die Fehler und wie werden sie korrigiert?*

Die Veränderungen an der Erwerbssequenz lassen sich vor allem dadurch erklären, dass Rekonstruktionsschritt 1 (Zustandsbeschreibung) und 2 (Übergangsrekonstruktion) abweichend ablaufen. Wenn die ersten beiden Schritte aber einmal korrekt bewältigt sind, wird das fragliche Ereignis bei Schritt 3 im Allgemeinen auch seltener, geringfügiger und weitestgehend symmetrisch fehldatiert. Dies zeigte sich daran, dass bei äquivalenter Sequenzrekonstruktion auch die Abweichungen der Dauerverteilung sehr gering waren (Tab. 5.7, sowie darin, dass übereinstimmend beschriebene Übergänge seltener fehldatiert wurden (Abb. 5.8)<sup>10</sup>.

Dies unterstreicht, wie wichtig es ist, dass in der Befragung die Episodenstruktur in den Schritten 1 und 2 korrekt rekonstruiert wird, besonders für Information über Arbeitslosigkeiten und Nichterwerbstätigkeiten. Diese werden dabei in ganzen „Episodenpaketen“ vergessen oder hinzugefügt, während abweichende Inzidenzen und Frequenzen von Erwerbstätigkeit eher dadurch hervorgerufen werden, dass Zeiträume in Erwerbstätigkeit durch abweichend rekonstruierte Übergänge in andere Episoden unterteilt oder Erwerbsepisoden gestreckt und gestaucht werden.

Der Erfolg der Edition zeigt sich deutlich in der besseren Übereinstimmung für fast alle Inhalte und hinsichtlich aller Fehleraspekte. Es gab keine Anhaltspunkte dafür, dass diese Verbesserungen durch Verschlechterungen an anderer Stelle „erkauft“ werden oder die Dateneingriffe eine spezifische Komponente von *Post-Survey Processing Error* einführen (Groves, 1989). Allerdings profitieren die Informationen, die auch schon bei der Erhebung inkonsistenter sind, auch von der Edition in geringerem Maß, nämlich die innerbetrieblichen Wechsel und die Verwechslung von Arbeitslosigkeiten und Nichterwerbstätigkeiten. Übergänge, die keine zentralen Veränderungen mit sich bringen, werden deswegen erstens weniger stabil rekonstruiert, und zweitens ruft es selten Widersprüche innerhalb der Angaben hervor, wenn sie eingeebnet werden. Ebenso können sich Arbeitslosigkeiten und Nichterwerbstätigkeiten in ihren Aktivitäten und Rollen ähneln, sodass das eine in das andere umgedeutet werden kann, ohne dass dies zu Widersprüchen in den Angaben führt. Ein eingehenderer Vergleich der edierten Angaben übersteigt Rahmen und Anspruch dieser Arbeit, könnte aber Aufschlüsse darüber geben, welche Editionsmaßnahmen am effizientesten sind und wo noch Optimierungsspielraum besteht. Hierzu sollten auch die edierte Version der ersten Befragung und die umfassenden Dokumentationen der Editionsarbeit herangezogen werden, bei der jede Veränderung im Datensatz transparent und nachvollziehbar dokumentiert wurde (Brückner, 1995; Wagner, 1989; Wagner & Visser, 2004).

*Welche Angabe ist die korrektere?*

Der Vergleich der Angaben aus einer Wiederholungsbefragung zeigt, wie sich autobiografische Rekonstruktionen derselben Sachverhalte über die Zeit hinweg verändert haben. In Ermangelung befragungsexterner Information über die tatsächlichen Erwerbsverläufe besteht leider keine Möglichkeit, diese Veränderung relativ zum tatsächlichen Sachverhalt zu betrachten. Die Frage, welche der beiden Angaben korrekter ist, muss also gestellt und erörtert werden.

<sup>10</sup> Fehldatierungen können allerdings zu deutlich verzerrter Information führen, wenn korrekt rekonstruierte Episoden dadurch aus dem Vergleichszeitraum hinausfallen bzw. irrtümlich hineinkommen.

Die Angaben aus der ersten Befragung haben aufgrund der kürzeren retrospektiven Distanz einen „Vertrauensvorschuss“, und tatsächlich deuten die Veränderungen auch auf einen allgemeinen gedächtnisbedingten Verlust von Information in der zweiten Befragung hin. Wenn wir allerdings Datengüte als die Angemessenheit begreifen, mit der ein sozialwissenschaftlicher Tatbestand in den Daten abgebildet wird (Rohwer & Pötter, 2002), gibt es auch Grund zur Annahme, dass an manchen Punkten die Angaben aus der zweiten Befragung angemessener sind. Gerade in turbulenten Zeiten können aus der längeren Rückschau Geschehnisse möglicherweise angemessener bewertet und eingeordnet werden. Besonders im Hinblick auf Arbeitslosigkeit ist durchaus vorstellbar, dass in der zweiten Befragung aufgrund der längeren Vertrautheit mit dem marktwirtschaftlichen System die Anwendung des Begriffs eher im Sinne der LV-Konzepte erfolgt. Auch die Reduktion des sozialen Stigmas der Arbeitslosigkeit nimmt den normativen Druck auf die Befragten, derartige Ereignisse aus ihrem Lebenslauf zu tilgen oder umzudeuten.

In der Einzelfallbetrachtung lassen sich in den reichhaltigen Detailinformationen Anhaltspunkte dafür finden, welche der beiden Sequenzen den Erwerbsverlauf angemessener abbildet. Wenn wir die Einzelfälle in Abbildung 5.5 betrachten, so haben Fall 2b und 2c offensichtlich in der zweiten Befragung Phasen der Arbeitslosigkeit einfach unterschlagen, und die Angaben aus der ersten Befragung sind angemessener. In anderen Fällen hingegen wird man auch durch den qualitativen Vergleich nicht klüger, beispielsweise im Fall 11c. Teilweise bietet auch eine der beiden Befragungen dem Datenanwender mehr Möglichkeiten, über zwiespältige Zustände je nach Erkenntnisinteresse zu entscheiden. In Fall 11b kann mit den Daten der zweiten Befragung die geringfügige Beschäftigung für 450 DM im Monat als Erwerbstätigkeit gewertet oder für den gesamten Zeitraum Nichterwerbstätigkeit eingetragen werden. Oder bei Fall 8a kann mit den Daten der ersten Befragung der Zeitraum mit Kurzarbeit Null nach Ermessen eingeordnet werden, während man mit den Daten der zweiten Befragung auf Arbeitslosigkeit festgelegt ist.

Gerade in Bezug auf die Rekonstruktion von Arbeitslosigkeit geben die Befunde Anlass, genauer darüber nachzudenken, *was* für Zeiträume hier eigentlich erfasst werden. Im nächsten Kapitel werde ich retrospektive Befragungsdaten zur Arbeitslosigkeit aus einer anderen Erhebungswelle der LV, der LV-West, mit den Aufzeichnungen der Bundesanstalt für Arbeit vergleichen. Dadurch werde ich einen genaueren Blick auf die Frage werfen können, was die Befragten unter Arbeitslosigkeit verstehen und ob sich Rekonstruktionen von Arbeitslosigkeiten zu früheren und späteren Zeitpunkten darin unterscheiden.

## Kapitel 6

### Arbeitslosigkeit in der Lebensgeschichte: Eine externe Validierung mit Daten des Arbeitsamts

#### 6.1 Fragestellung und Vorgehen

Bei der retrospektiven Erhebung von Erwerbsgeschichten in Befragungen hat sich gezeigt, dass Arbeitslosigkeit von allen Zuständen der Arbeitsmarktbeteiligung am instabilsten rekonstruiert wird. Die valide bzw. angemessene Abbildung von Arbeitslosigkeit stellt also eine anspruchsvolle Herausforderung dar. Wiederholungsbefragungen und externe Validierungsstudien zeigen eine Neigung zum Vergessen und Unterschlagen von Arbeitslosigkeiten zumindest für stabile Arbeitsmärkte (siehe Kap. 5). Vor allem werden kurze und kürzeste Episoden unterschlagen, und es gibt Hinweise darauf, dass die Angaben mit längerer retrospektiver Distanz schlechter werden (Dex & McCulloch, 1997; Elias, 1997; Mathiowetz & Duncan, 1988; Paull, 2002).

Da diese Studien sich meistens nur auf einige wenige Jahre beziehen, ist wenig darüber bekannt, wie Arbeitslosigkeiten über lange Zeiträume im Rahmen einer Lebensgeschichte repräsentiert und rekonstruiert werden. Außerdem gibt die Literatur nahezu ausschließlich darüber Auskunft, ob und in welchem Umfang Phasen, die eigentlich als Arbeitslosigkeit angegeben werden sollten, nicht angegeben werden, während umgekehrt wenig darüber bekannt ist, in welchem Umfang und unter welchen Bedingungen für Zeiträume zu Unrecht Arbeitslosigkeit angegeben wird.

Diese Fragen werde ich mit einem Vergleich der retrospektiven Angaben zur Erwerbsgeschichte aus der Befragung *Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland (LV-West)* mit der Leistungsempfängerdatei (LED) der Bundesanstalt für Arbeit (BA) untersuchen. Diese prozessproduzierten Daten werden vom Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB) bereitgestellt und können im Rahmen einer Validierungsstudie auf Personenebene mit den Befragtenangaben verknüpft werden. Dadurch können die vollständigen Erwerbskarrieren der Befragten bis zum Befragungszeitpunkt verglichen werden, die einen Zeitraum von bis zu 18 Jahren abdecken (siehe Abb. 6.1). Es handelt sich dabei um einen „vollständigen Registerdatenvergleich“ (Groves, 1989; Marquis, 1978), bei dem sowohl nichtberichtete Leistungsbezugszeiträume als auch berichtete Arbeitslosigkeiten ohne entsprechenden Leistungsbezug betrachtet werden können.

Mit dem Vergleich der Datensätze möchte ich folgenden Fragen nachgehen:

- (a) Wie werden Leistungsbezugszeiträume in den Befragungsdaten abgebildet? Welche Arten von Abweichungen treten auf, welche Episoden sind davon besonders betroffen, und was sagt uns das über die Repräsentation und Rekonstruktion von Arbeitslosigkeit innerhalb der Erwerbsbiografie?
- (b) Wie entsprechen die in der Befragung berichteten Arbeitslosigkeiten den Leistungsbezugszeiträumen? In welchem Umfang werden Zeiträume als Arbeitslosigkeiten rekonstruiert,

Abbildung 6.1: Vergleichszeitraum für den Vergleich von LED und LV

	Befragung												1. Leistungsbezugszeitraum 1964er Kohorte		1. Leistungsbezugszeitraum 1971er Kohorte		1. Leistungsbezugszeitraum 1964er Kohorte						
	1998	97	96	95	94	93	92	91	90	89	88	87	86	85	84	83	82	81	80	79	78	77	
Jahr																							
Alter der 1964er Kohorte	34	33	32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	
Alter der 1971er Kohorte	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8	7	6	
Retrospektive Distanz (Jahre)	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	
Arbeitslosenquote (in %)	10,5	11,0	10,1	9,3	9,2	8,2	6,6	7,2	7,9	8,7	8,9	9,0	9,3	9,1	9,1	7,5	5,5	3,8					

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsenmpfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen; Statistisches Jahrbuch für die Bundesrepublik Deutschland, 2000.

in denen der Befragte nachweislich keine Leistungen bezog, welche Episoden sind besonders betroffen, und was sagt uns das über die Repräsentation und Rekonstruktion von Arbeitslosigkeit innerhalb der Erwerbsbiografie?

- (c) Weichen die Registereinträge und die Befragungsdaten für länger zurückliegende Zeiträume stärker voneinander ab, oder fallen die Abweichungen für länger zurückliegende Zeiträume anders aus?

Die Befragungsdaten der LV-West enthalten eine Fülle von Detailinformationen über die Arbeitslosigkeitsepisoden selbst und über die Episoden davor und danach. Außerdem lässt sich verfolgen, an welchen Stellen des Interviews welche Angaben gemacht wurden und wie sie in die Datenbank gelangt sind. Diese Kontextinformation kann herangezogen werden, um Abweichungen zwischen den Datensätzen aus dem Datenkonstruktionsprozess heraus zu erklären. Ich werde die uneditierten Daten der LV-West verwenden. Zwar kommen die edierten Daten der Realität vermutlich näher, aber nur die uneditierten Daten stellen die originären Produkte der Erinnerungsarbeit der Befragten in der vom Projekt konzipierten und gestalteten Abrufsituation dar und geben somit eher Aufschluss über die hier ablaufenden Prozesse als die edierten Daten.

Im ersten Abschnitt dieses Kapitels werde ich die beiden Datensätze, ihre Datengenerationsprozesse und ihre Definition von Arbeitslosigkeit beschreiben. Ich erörtere die Frage der Vergleichbarkeit der Datensätze und der Validität der Abbildung von Arbeitslosigkeit in prozessproduzierten und Befragungsdaten und leite daraus Vermutungen über Ausmaß und Art der Abweichungen in beide Vergleichsrichtungen ab.

Im zweiten Abschnitt werde ich auf Einzelfallebene qualitativ verschiedene Arten von Inkonsistenzen identifizieren, die zwischen den Angaben auftreten und teilweise als Folge der kognitiven Prozesse in der LV-West betrachtet werden können. Im dritten Abschnitt werde ich das Ausmaß der Abweichung quantitativ auf der Ebene von Monaten und Episoden beschreiben und mit der Länge der Episoden sowie mit der retrospektiven Distanz in Verbindung bringen. Nach einer Zusammenfassung der Befunde und ihrem Vergleich mit den bisherigen werde ich zuletzt besprechen, was die Befunde über die Rekonstruktion von Arbeitslosigkeit aussagen, und Implikationen diskutieren, die sich für die Erhebung und Verarbeitung von retrospektiven Angaben zu Arbeitslosigkeit ergeben.

## 6.2 Arbeitslosigkeit: Konzept, Definition, Erhebung

Bei Arbeitslosigkeit handelt es sich um einen sozialwissenschaftlichen Tatbestand, der in der Arbeitsmarktforschung bei einer Person häufig durch einen der folgenden drei Indikatoren festgestellt wird:

- ob eine Person sich selbst als arbeitslos bezeichnet,
- Leistungsbezugs- bzw. Antragstellerstatus einer Person (*claimant count*),
- eine von der *International Labor Organisation* (ILO) empfohlene tätigkeitsbasierte Definition, die eine Person dann als arbeitslos führt, wenn sie zum Befragungszeitpunkt keine Arbeit hat, aber arbeiten will, in den vergangenen vier Wochen aktiv nach Arbeit gesucht hat und innerhalb der nächsten zwei Wochen eine Erwerbstätigkeit aufnehmen kann oder wenn sie bereits eine Stelle gefunden hat, die sie innerhalb der nächsten zwei Wochen antreten wird.



Die LED enthält Informationen über den Antragstellerstatus zu verschiedenen Zeiten, die LV-West darüber, für welche Zeiträume eine Person sich in der Rückschau selbst als arbeitslos bezeichnet. Untersuchungen, in denen Arbeitslosigkeit für den Zeitpunkt der Befragung auf mehrere Weisen erhoben wurden, zeigen, dass sich die Indikatoren nicht genau decken und die Arbeitslosenquoten, die nach diesen Definitionen für eine Population gemessen werden, sich unterscheiden (Dex & McCulloch, 1997; Gallie & Vogler, 1994; McGinnity, 2001; Norwood & Tanur, 1994). Es lässt sich vermuten, dass auch retrospektiv nicht alle Leistungsbezugszeiträume als Arbeitslosigkeiten rekonstruiert werden; diese Art der Inkonsistenz möchte ich „Typ A“ nennen. Umgekehrt ist anzunehmen, dass über den Antragstellerstatus einige Zeiträume nicht erfasst werden, in denen sich die Befragten selbst als arbeitslos einstufen. Diese Art der Inkonsistenz möchte ich „Typ B“ nennen.

### 6.2.1 Arbeitslosigkeit in der LV-West

#### *Konzept und Definition*

Im Rahmen der *Deutschen Lebensverlaufstudie* (LV), Teilprojekt LV-West, des Max-Planck-Instituts für Bildungsforschung, Berlin, unter der Leitung von Prof. K. U. Mayer wurde in den Jahren 1998 und 1999 in Zusammenarbeit mit dem Institut für Arbeitsmarkt- und Berufsforschung der Bundesanstalt für Arbeit (IAB)<sup>1</sup> eine repräsentative Stichprobe von 2.900 Mitgliedern der westdeutschen Geburtskohorten 1964 und 1971, inklusive deutschsprechender Ausländer, zu ihren Erwerbsverläufen retrospektiv per CATI oder CAPI<sup>2</sup> befragt. Information über Zeiträume der Arbeitslosigkeit wird im Rahmen der Erwerbsgeschichte erhoben, die nach dem Ende der allgemein bildenden Schule bzw. nach dem Ende der Erstausbildung beginnt. Der Erwerbsverlauf besteht dabei aus Erwerbstätigkeiten (als der „Standardaktivität“) und aus mit verschiedenen Aktivitäten ausgefüllten Lücken, wovon eine Arbeitslosigkeit ist. Der Begriff „Arbeitslosigkeit“ wurde nicht genauer definiert, weder für den Befragten noch für den Interviewer; der soziale Tatbestand der Arbeitslosigkeit wird in der LV-West also über die *retrospektive Selbstdefinition* des Befragten erhoben.

#### *Datengenerierung und Datenfehlergenerierung*

Arbeitslosigkeiten werden in der LV-West als Lücken in der Erwerbsgeschichte erhoben. Die Befragten definierten zunächst ihren Zustand nach Abschluss der allgemein bildenden Schule bzw. Ausbildung, rekonstruierten dann den nächsten Übergang und datierten Start und Ende der Episode. Außerdem gaben sie an, ob sie im Anschluss daran sofort oder nach einer Unterbrechung eine weitere Erwerbstätigkeit aufgenommen oder sich gänzlich aus der aktiven Erwerbsarbeit verabschiedet hätten. Bei zeitweiser Unterbrechung wird vom CATI-Programm automatisch eine „Lücke“ für den Befragten gespeichert. Nach Abschluss der Erhebung aller Erwerbsepisoden werden die Befragten zu ihren Aktivitäten während dieser unabgedeckten Zeiträume befragt. Die Befragten beschreiben ihre Aktivität am Beginn der Lücke (als

<sup>1</sup> Dort wurde die Studie von Hans Dietrich und Stefan Bender betreut und durch Mittel des europäischen Sozialfonds gefördert.

<sup>2</sup> *Computer Assisted Telephone Interview* und *Computer Assisted Personal Interview*.

Abbildung 6.2: Datengenerierungsprozess bei der Erhebung von Arbeitslosigkeiten in der LV-West

Schritt 1	<i>Übergang rekonstruieren:</i> Veränderungen in den Erwerbs- und arbeitsmarktbezogenen Aktivitäten, Rollen, Orte und Personen rekonstruieren
Schritt 2	<i>Episode benennen:</i> Erwerbs- und arbeitsmarktbezogene Aktivitäten, Rollen, Orte und Personen rekonstruieren und einer Kategorie „arbeitslos“ zuordnen
Schritt 3	<i>Übergang datieren:</i> Monat und Jahr des Episodenanfangs und -endes rekonstruieren

Arbeitslosigkeit oder eine von acht Nichterwerbstätigkeiten<sup>3</sup>), rekonstruieren das Ende dieser Aktivität und datieren es. Ist damit noch nicht die ganze gespeicherte Lücke abgedeckt, wird nach der nächsten Aktivität innerhalb dieser Lücke gefragt; andernfalls folgt die nächste Lücke.

Eine Arbeitslosigkeitsepisode kommt also durch einen dreistufigen Rekonstruktionsprozess zu Stande (siehe Abb. 6.2). Als erster Schritt muss ein Zeitraum als distinkt von den angrenzenden Erwerbstätigkeiten oder Nichterwerbstätigkeiten rekonstruiert werden, das heißt, eine Veränderung in den arbeitsmarkt- und erwerbsbezogenen Rollen, Aktivitäten usw. muss rekonstruiert werden. Bei Schritt 2 muss für diesen Zustand als angemessenste Beschreibung „Arbeitslosigkeit“ gewählt werden<sup>4</sup>, und bei Schritt 3 müssen zuletzt Anfang und Ende datiert werden (bzw. muss nur das Ende datiert werden, da das Anfangsdatum sich bereits aus der Angabe des Enddatums der vorangehenden Episode ergeben hat).

Abweichungen der rekonstruierten Arbeitslosigkeitsepisoden von den in der LED erfassten Leistungsbezugszeiträumen können auf allen drei Stufen verursacht werden:

- wenn am Beginn eines Leistungsbezugszeitraums kein Wechsel in den Rollen und Aktivitäten wahrgenommen und deshalb kein Übergang in Arbeitslosigkeit rekonstruiert wird (Typ A), oder umgekehrt, wenn ein solcher Wechsel in eine Arbeitslosigkeit wahrgenommen wird, ohne dass ein Leistungsbezugszeitraum vorliegt (Typ B),
- wenn Übergänge übereinstimmend rekonstruiert werden, die fragliche Periode aber nicht als Arbeitslosigkeit wahrgenommen und beschrieben wird (Typ A),
- wenn die Episoden und Übergänge zwar prinzipiell übereinstimmend mit den Zeiträumen des Leistungsbezugs rekonstruiert werden, aber durch Fehler beim Datieren nicht denselben Zeitraum abdecken (Typ A und Typ B).

<sup>3</sup> Es handelte sich hierbei um: in beruflicher Ausbildung – Berufsschule; hauptberuflich erwerbstätig über 15 Std. die Woche (auch Arbeitsbeschaffungsmaßnahme); im Erziehungsurlaub – Mutterschutz; geringfügig erwerbstätig – Gelegenheitsjobs, Studentenjobs; im Wehrdienst; im Zivildienst; krank – in Rehabilitation – Kur; in einer allgemein bildenden Schule.

<sup>4</sup> In diesem Zusammenhang ist wichtig zu bedenken, dass der Befragte, der einmal eine Lücke in seiner Erwerbsgeschichte rekonstruiert hat, unter *allen angebotenen* Aktivitätsbeschreibungen die *passendste* für diesen Zeitraum auswählt. Schwarz und Hippler (1991) konnten vielfach zeigen, dass die Interpretation dessen, was und wie berichtet werden soll, auch davon abhängt, welche Antwortkategorien zur Verfügung stehen. In den ostdeutschen Teilstudien der LV-Ost beispielsweise war unter den Nichterwerbstätigkeiten der Zustand „Hausfrau“; in dieser LV-West hingegen gibt es nur die Kategorie „anderes (z.B. Hausfrau/-mann, Kindererziehung, Erholung, Sonstiges)“. Hausfrauen- und Arbeitslosigkeitszustände ähneln einander teilweise von den Aktivitäten her, wodurch Zuordnungsschwierigkeiten entstehen, die vielleicht eher in Richtung Hausfrau aufgelöst werden, wenn dies als explizite vollwertige Kategorie angeboten wird.

In Kapitel 5 haben wir gesehen, dass bei der Rekonstruktion von Erwerbsverläufen Fehler beim Datieren eine vergleichsweise geringe Rolle spielen. Lediglich durch ungünstiges Zusammenfallen mit den Grenzen des Zeitfensters hatten sich hier größere und systematische Abweichungen zwischen den beiden Befragungen ergeben. Ausschlaggebend ist also, wie Übergänge retrospektiv rekonstruiert werden, wie die rekonstruierten Episoden beschrieben werden und ob dies in Übereinstimmung mit den Zeiträumen des Leistungsbezugs geschieht.

### 6.2.2 Arbeitslosigkeit in der LED

In der LED der BA werden seit 1975 alle Zeiträume erfasst, in denen Personen Lohnersatzleistungen von der Bundesanstalt für Arbeit bezogen haben. Nach Bick und Müller (1984) entstehen prozessproduzierte Daten, indem die Behörde Menschen Information im Rahmen von Verwaltungskontakten abverlangt. Dies geschieht vermittelt durch Formulare und andere Aufzeichnungsvorschriften. Dieser Prozess wird beeinflusst durch den *Aktenproduzenten*, die *Verwaltungstheorie* (oder Verwaltungsnormen) und den *Klienten* (vgl. Abschnitt 3.3.5).

Darüber, wie der Aktenproduzent und der Klient im Verlauf des Datengenerationsprozesses Fehler in der Abbildung der Bezugszeiträume verursachen, ist – wie bei prozessproduzierten Daten häufig der Fall – nichts bekannt. Über die Verwaltungsnormen, also die Regelungen, welche Information von welchen Personen in welcher Form in welchen Zeitabständen auf welchem Wege erhoben werden *soll*, informieren Bender, Hilzendege, Rohwer und Rudolph (1996) und der Leistungs-Informationen-Service des Arbeitsamts (<http://www.arbeitsamt.de/hst/services/lis/index.html>, Version von August 2003). Ziel der LED ist die Erfassung der Zeiträume, in denen die betreffenden Personen von der BA *Arbeitslosengeld*, *Arbeitslosenhilfe* oder *Unterhaltsgeld* erhielten. Für den Bezug dieser Leistungen werden derzeit folgende Voraussetzungen genannt:

- Für den Bezug von *Arbeitslosengeld* müssen sich die Personen persönlich beim Amt als arbeitslos melden. Im Allgemeinen muss in den vergangenen drei Jahren mindestens 360 Tage lang ein sozialversicherungspflichtiges Beschäftigungsverhältnis bestanden haben, 180 Tage für Wehr- und Ersatzdienstleistende sowie Saisonarbeiter. Antragsteller müssen dem deutschen Arbeitsmarkt prinzipiell zur Verfügung stehen, das heißt, sie dürfen nicht längerfristig erkrankt oder in Rehabilitation sein, müssen sich im Lande aufhalten und dürfen keiner Beschäftigung nachgehen, die über den Umfang von 15 Stunden hinausgeht. Für Personen mit einer anderen als der deutschen Staatsbürgerschaft müssen weitere Voraussetzungen erfüllt sein.
- *Arbeitslosenhilfe* wird, bei Ablauf oder Nichtbestehen des Anspruchs auf Arbeitslosengeld, ebenfalls nur auf persönlichen Antrag und nur bei Vorliegen weiterer Voraussetzungen gezahlt, wenn der Lebensunterhalt des Antragstellers auf keinem anderen Wege bestritten werden kann.
- *Unterhaltsgeld* wird ebenfalls auf persönlichen Antrag hin gewährt, um Arbeitslosen oder von Arbeitslosigkeit bedrohten Personen eine berufliche Weiterbildung zu ermöglichen. Um Unterhaltsgeld beziehen zu können, müssen in der Regel ebenfalls 360 Tage versicherungspflichtige Beschäftigung innerhalb der letzten drei Jahre bestanden haben, oder die Personen müssen Anspruch auf Arbeitslosengeld oder Arbeitslosenhilfe haben.

Auch bei bestehender Bezugsberechtigung können für festgelegte Sperrzeiten die Zahlungen ausgesetzt werden, wenn der Bezugsberechtigte seine Arbeitslosigkeiten durch eigene Kündigung oder Fehlverhalten am Arbeitsplatz ohne Not herbeigeführt hat, einer Beschäftigung über 15 Stunden die Woche im Bezugszeitraum nachgeht, zumutbare Stellenangebote ablehnt oder sich bei der Vermittlung unkooperativ verhält. Diese Sperrzeiten dauern in der Regel zwölf Wochen, in Einzelfällen auch drei oder sechs Wochen, und werden in der LED nicht als Leistungsbezugszeiträume geführt.

### 6.2.3 Vergleichbarkeit und Validität

In Abschnitt 3.3.5 waren folgende Voraussetzungen für eine Zusammenführung und Validierung von Befragungsdaten mit und an Registerdaten genannt worden: Vergleichbarkeit der Bezugsseinheiten, der Bezugszeiträume, der Informationen und der Datenformate.

LED und LV-West enthalten Information über drei *Bezugsseinheiten*: Personen, Episoden und Zeiteinheiten. 636 Personen können in beiden Datensätzen einander eindeutig zugeordnet werden, da die Teilnehmer ihre Sozialversicherungsnummer angegeben hatten, anhand derer ihre Einträge in der LED identifiziert werden können. Die Zeiteinheiten (z.B. Monate oder Jahre) können einander ebenso eindeutig zugeordnet werden, da beide Datensätze den julianischen Kalender zur Datierung von Zeiträumen von Leistungsbezug bzw. Arbeitslosigkeit verwenden. Die Episoden hingegen stellen eine komplexe Herausforderung dar: Aufgrund der Abweichungen zwischen den beiden Datensätzen ist nicht zu erwarten, dass LED und LV-West identische oder einander unproblematisch zuordenbare Episoden enthalten. Die „technische“ Frage der Vergleichbarkeit der Arbeitslosigkeitsepisoden wird hier also identisch mit der methodischen Frage der Validität der erhobenen Arbeitslosigkeitsepisoden.

Hinsichtlich der *Bezugszeiträume* sind LED und LV-West weitestgehend vergleichbar. In der LV-West werden die Erwerbsgeschichten beginnend mit dem Abschluss der allgemein bildenden Schule, die die Teilnehmer frühestens Ende der 1970er Jahre beendeten, bis zum Datum des Interviews 1998 oder 1999 erhoben. Die LED umfasst sämtliche Leistungsbezugszeiträume ab 1975 (wobei erst die Angaben ab etwa 1980 in auswertbarer Form zur Verfügung stehen). Abbildung 6.1 zeigt für beide Kohorten getrennt, innerhalb welcher Zeiträume Angaben zur Arbeitslosigkeit in LED und LV vorliegen. Der erste Registereintrag und die erste Angabe liegen für die jüngere Kohorte natürlich einige Jahre später vor. Durch das „Abschneiden“ der LED-Daten am Interviewzeitpunkt werden die Zeiträume vergleichbar.

Die *Datenstruktur* der Datensätze ist unterschiedlich: In der LV-West sind alle ununterbrochenen Episoden von Arbeitslosigkeit in den Datenzeilen repräsentiert, die durch eine Fallnummer der Person, die sie berichtet hat, zugewiesen sind („Spelldatensatz“). In der LED stehen in den Datenbankzeilen meist ebenfalls zusammenhängende Leistungsbezugsepisoden einer durch eine Fallnummer zuordenbaren Person; diese Episoden können dabei aber in der Regel nur ein Jahr lang sein. Wenn eine Person länger als ein Jahr Leistungen bezieht, finden sich entsprechend mehr zeitlich lückenlos aufeinander folgende Episoden in aufeinander folgenden Datenzeilen. Vergleichbarkeit lässt sich durch das Aneinanderfügen solcher Arbeitslosigkeitsepisoden in der LED zu einer zusammenhängenden Episode herstellen.

Hinsichtlich der Vergleichbarkeit der *Information* müssen vor allem zwei Aspekte diskutiert werden: Zum einen ist die *Definition* von Arbeitslosigkeit in der LV-West nicht identisch mit der Definition von Leistungsbezugszeiträumen. Zum anderen werden die *Start- und Enddaten in den beiden Datensätzen unterschiedlich erfasst*.

Was die Vergleichbarkeit der erhobenen Konzepte von Arbeitslosigkeit angeht, so berührt dies die zentrale methodologische und theoretische Frage dieses Kapitels: Wie wird Arbeitslosigkeit im autobiografischen Gedächtnis repräsentiert und im Kontext der standardisierten Befragung retrospektiv rekonstruiert und wie entsprechen diese biografischen Rekonstruktionen den Leistungsbezugszeiträumen? Dieser Frage sind die folgenden Abschnitte sowie die empirischen Analysen gewidmet.

Was die Vergleichbarkeit der Datumsangaben angeht, so werden Start- und Enddatum in der LV-West nur als Monat und Jahr erhoben, in der LED hingegen bis auf den Tag genau. Dadurch können Episoden in der LED auch nur einen Tag lang sein und innerhalb eines Monats mehrere Episoden auftreten, während in der LV-West die Mindestdauer jeder Arbeitslosigkeit einen Monat beträgt – selbst wenn die tatsächliche Arbeitslosigkeit nur drei Wochen oder noch weniger dauerte. Um die Daten dennoch vergleichbar zu machen, sind einige Entscheidungen nötig, die auf der Grundlage des Wissens um die Datenkonstruktionsprozesse und das Ziel des Vergleichs getroffen werden müssen. Mein diesbezügliches Vorgehen wird in Abschnitt 6.4 beschrieben und begründet. In jedem Fall kommt es durch die Anpassung zu Anschlussproblemen und Rundungsunschärfen.

#### 6.2.4 Annahmen über Abweichungen

##### *Welche Abweichungen treten auf?*

Da in der LV-West der Begriff „arbeitslos“ nicht näher definiert ist, rekonstruiert der Befragte Arbeitslosigkeit gemäß seiner persönlichen Interpretation dieses Begriffs. Wenn diese Interpretation nicht für alle Zeiträume gilt, in denen Leistungen bezogen wurden, oder umgekehrt auch für Zeiträume gilt, in denen keine Leistungen bezogen wurden, kann es hier zu Abweichungen kommen.

Die Erwerbsgeschichte stellt eine zentrale thematische Achse im autobiografischen Gedächtnis dar. Arbeitslosigkeiten von mindestens einem Monat Dauer, wie die LV-West sie erhebt, sind am ehesten als Lebensabschnitte im Sinne von Conway und Pleydell-Pearce (2000) repräsentiert. Diejenigen Leistungsbezugszeiträume, die den zum Zeitpunkt der Rekonstruktion als Arbeitslosigkeit repräsentierten Episoden entsprechen, haben die größte Chance, in den Befragungsdaten übereinstimmend genannt zu werden.

Wie in Abschnitt 2.3 beschrieben, bilden sich Lebensabschnitte innerhalb eines thematischen Bereichs dadurch, dass die konstitutiven sozialen Rollen, Aktivitäten, Orte und beteiligten Personen für einen ausgedehnten Zeitraum gleich bleiben. Übergänge entstehen dort, wo diese sich auf persönlich bedeutsame und folgenschwere Weise ändern. Neben der bloßen Abwesenheit von Erwerbsarbeit (wie sie auch die acht Zustände der Nichterwerbstätigkeit kennzeichnen), werden Leistungsbezugszeiträume durch eine Reihe von erwerbs- und arbeitsmarktbezogenen Aktivitäten als Arbeitslosigkeit ausgewiesen: Der Anwärter muss das Arbeitsamt aktiv aufsuchen, sich selbst zum Arbeitslosen erklären und den Nachweis der Berechtigung erbringen. Nach Bewilligung von Arbeitslosengeld oder -hilfe steht er

Abbildung 6.3: Hypothesen im Überblick

---

<i>Inkonsistenzen vom Typ A:</i>
Kurze Leistungsbezugszeiträume ...
Leistungsbezugszeiträume, während derer andere Aktivitäten im Vordergrund standen ...
Länger zurückliegende Leistungsbezugszeiträume ...
Leistungsbezugszeiträume zu Beginn der Erwerbskarriere ...
Leistungsbezugszeiträume in der unmittelbaren Vergangenheit ...
... werden seltener als Arbeitslosigkeiten rekonstruiert
<i>Inkonsistenzen vom Typ B:</i>
Kurze Arbeitslosigkeiten ...
Länger zurückliegende Arbeitslosigkeiten ...
Arbeitslosigkeiten zu Beginn der Erwerbskarriere ...
... gehen häufig nicht oder nur teilweise nicht mit Leistungsbezug einher

---

in ständigem Kontakt mit dem Arbeitsamt, sowohl bei der Entgegennahme der Zahlungen als auch beim Nachweis der aktiven Bemühung um Einstellung bzw. bei der Inanspruchnahme der Vermittlungshilfe des Arbeitsamts. Diese Aktivitäten unterscheiden sich deutlich von denen, die Erwerbstätigkeit konstituieren, und etwas weniger deutlich auch von denen, die den verschiedenen Nichterwerbstätigkeiten eigen sind (Ähnlichkeiten könnten besonders bei Personen bestehen, die Sozialhilfe oder andere staatliche Unterstützungsleistungen beziehen). Leistungsbezugszeiträume sollten also von allen möglichen Zuständen ohne Erwerbsarbeit die besten Voraussetzungen haben, auch als Arbeitslosigkeit repräsentiert zu sein.

Allerdings wurde gezeigt, dass Arbeitslosigkeit innerhalb der Erwerbsbiografie nicht als vollwertiger Zustand mit eigenständigen erwerbs- und arbeitsmarktbezogenen Rollen und Aktivitäten konstruiert wird, sondern im Kontext der Erwerbsarbeit als deren Negation, deren unfreiwilliger und unerwünschter Verlust und Abwesenheit (Elias, 1997; Mutz, 1996). Als solche *Non-States* (Elias, 1997) sind Arbeitslosigkeitsphasen im Rahmen der ziel- und entscheidungsbasierten Logik der Lebensgeschichte stets in besonderem Maße begründungsbedürftig: Warum wurde gerade diese Person zu gerade diesem Zeitpunkt arbeitslos? Auch die eindeutig negative, sozial stigmatisierte Erfahrung von Arbeitslosigkeit führt dazu, dass Umdeutungen in andere Zustände sehr bereitwillig vorgenommen werden, wenn dies möglich ist.

Auf dieser Grundlage ist zu erwarten, dass relativ kurze Zeiträume des Leistungsbezugs häufig nicht als Lebensabschnitte repräsentiert sind (Inkonsistenzen vom Typ A). Das Gleiche gilt für Arbeitslosigkeitsphasen, deren Ende bereits bei Antragstellung abzusehen war – etwa durch das Angebot einer Arbeitsstelle, oder der Gewissheit, schnell wieder unterzukommen, etwa bei Vertretern einer gesuchten Berufsgruppe. Wie bereits angedeutet, ist der Anreiz in Deutschland, sich auch in solchen Fällen arbeitslos zu melden, im Vergleich zu anderen Ländern sehr hoch, da bei Erfüllung der Voraussetzungen vergleichsweise großzügig bemessene Zahlungen und weit reichender Versicherungsschutz gewährt werden. Weiterhin anfällig sind Zeiträume, in denen trotz Leistungsberechtigung und -bezug für den Befragten andere Rollen und Aktivitäten im Vordergrund standen und etwa trotz Leistungsbezug einer Erwerbsarbeit nachgegangen wurde (im Rahmen der erlaubten 15 Stunden pro Woche oder „schwarz“ darüber hinaus), wenn in diesem Zeitraum auch Hausarbeit, Kindererziehung oder Angehö-

riegenpflege verrichtet wurde oder der Leistungsbezieher selbst lange krank oder in Rehabilitation war<sup>5</sup>.

Umgekehrt nehme ich an, dass auch Zeiträume in der Lebensgeschichte als Arbeitslosigkeiten repräsentiert sind, in deren Verlauf aber keinerlei Leistungen bezogen wurden (Inkonsistenzen vom Typ B). Zum einen betrifft dies Personen und Episoden, bei denen die Voraussetzungen für Leistungsbezug nicht gegeben sind und der entsprechende Antrag deshalb nicht gestellt oder abgelehnt wurde<sup>6</sup>. Dies ist zum Beispiel der Fall bei in Deutschland lebenden Ausländern, bei jungen Menschen, die noch nicht lange genug sozialversicherungspflichtig beschäftigt waren, oder bei Frauen, die nach ausgedehnten Familienphasen in Erwerbstätigkeit zurückkehren möchten. Weiterhin sind die bereits erwähnten Sperrzeiten in der LED nicht enthalten, während die Zahlungsunterbrechung für die gesperrte Person vermutlich selten einen Übergang in einen anderen Zustand darstellt.

#### *Länge der retrospektiven Distanz*

Welche Rolle spielt die Länge der Distanz zu einem Ereignis für die Übereinstimmung zwischen LED und LV-West; macht es einen deutlichen Unterschied, ob Arbeitslosigkeiten rekonstruiert werden sollen, die zwei oder zehn Jahre zurückliegen? Für Arbeitslosigkeit wurden für Zeiträume von bis zu sechs Jahren folgende Vergessenskurven angenommen oder nachgewiesen:

- linear oder zumindest kontinuierlich zunehmendes Vergessen,
- bis zu drei Jahren vor der Befragung konstant gutes Erinnern, danach zunehmendes Vergessen (Elias, 1997),
- schlechteres Erinnern an Erwerbs- und Arbeitsmarktaktivitäten jüngerer Personen in der beruflichen Einstiegsphase (Paull, 2002),
- schlechtere Angaben für die Zeit kurz vor der Befragung, weil Arbeitslosigkeit als stigmatisiertes Ereignis ungern zugegeben wird.

Beim ersten Schritt der Rekonstruktion kann es sein, dass für länger zurückliegende Zeiträume der Beginn eines Leistungsbezugszeitraums seltener auch als Wechsel in einen neuen Lebensabschnitt rekonstruiert wird, da die Erinnerung auch mangels häufiger Übung verblasst, die Abruf-Hinweisreize weniger werden und auch die Dramatik und Folgeschwere aus der längeren Distanz relativiert werden. Dadurch müssten Inkonsistenzen vom Typ A für weiter zurückliegende Zeiträume häufiger auftreten.

Diese Distanz eines Leistungsbezugszeitraums zur Gegenwart steht auch hier wieder für unterschiedliche Rekonstruktionsbedingungen (siehe Abschnitte 2.6.3 und 3.4.3). Vor allem gehen längere retrospektive Distanzen mit früheren Phasen der Erwerbskarriere einher, sodass das Alter beim Erleben eines Leistungsbezugs relevant wird. Phasen des Leistungsbezugs zu Beginn der Erwerbsgeschichte – etwa zwischen Ausbildung, Wehrdienst und Ersterwerbstätigkeit oder zwischen Einstiegs- und Erwerbstätigkeiten – sind im Vergleich zu Arbeitslosigkeit später im Lebenslauf verbreitet, erwartet, kurz und weniger dramatisch und bruchhaft. Solche „Ein-

<sup>5</sup> In diesem Fall bleibt das Arbeitsamt bis zu sechs Wochen zuständig, danach wird an andere Träger, etwa Rentenkassen oder Sozialamt, übergeben.

<sup>6</sup> Natürlich ist es denkbar, dass auch für bestimmte Zeiträume kein Antrag gestellt wird, obwohl die Voraussetzungen für Leistungsbezug gegeben wären, besonders für kurze Episoden mit abzusehendem Ende, wenn der Betroffene keinen finanziellen Mangel leidet und deshalb den Aufwand scheut.

stiegsarbeitslosigkeiten“ sind deshalb mit einer höheren Wahrscheinlichkeit nicht als Lebensabschnitt in Arbeitslosigkeiten repräsentiert, und Inkonsistenzen vom Typ A werden für weiter zurückliegende Zeiträume häufiger auftreten.

Umgekehrt erwarte ich für den Beginn der Erwerbsgeschichte auch ein höheres Risiko, Arbeitslosigkeitsepisoden zu rekonstruieren, die nicht mit Leistungsbezug einhergehen, also vermehrt Inkonsistenzen vom Typ B. Am Übergang zwischen Schule, Ausbildung, Wehrdienst oder erster Erwerbstätigkeit gibt es vergleichsweise viele Zeiträume, in denen die Befragten sich als arbeitslos betrachten können, während derer aber aufgrund fehlender Anwartschaftszeiten keine Leistungen beantragt oder bewilligt wurden.

Eine schlechtere Übereinstimmung vom Typ A ist auch für die Zeiträume kurz vor der Befragung zu erwarten. Hier werden Leistungsbezugszeiträume seltener als Arbeitslosigkeiten rekonstruiert, wenn die Befragten ungern zugeben, arbeitslos zu sein. Dabei handelt es sich nicht um einen Gedächtnisfehler, sondern um absichtliches Verschweigen im Sinne einer sozialen Erwünschtheit. In Mathiowetz und Duncan (1988, S. 224, Abb.) wird zum Beispiel sichtbar, dass Befragtenangaben zu Arbeitslosigkeiten gerade für die Monate direkt vor der Befragung mit den Registerdaten besonders selten übereinstimmen. Allerdings gehen die Autoren nicht näher darauf ein und geben auch nicht an, ob es sich, wie bei einem Effekt sozialer Erwünschtheit zu erwarten, vor allem um Inkonsistenzen durch unterschlagene Registerinträge handelt.

In den folgenden Absätzen werde ich diese Annahmen durch den Vergleich der LED mit der LV-West empirisch überprüfen. Da das Wissen über die Rekonstruktion von Arbeitslosigkeiten über die gesamte Erwerbskarriere und Zeiträume von fast 20 Jahren noch begrenzt ist, werde ich zuerst mit einer ausgedehnten Einzelfallbetrachtung phänomenologisch klären, auf welche Weise sich die Angaben zu Arbeitslosigkeit zwischen den beiden Datensätzen unterscheiden und wie sich dies im Rahmen des oben beschriebenen dreistufigen Rekonstruktionsprozesses bzw. der unterschiedlichen Datenkonstruktionsprozesse begründen lässt. Danach werde ich den Grad der Übereinstimmung quantitativ beschreiben und mit der Länge der retrospektiven Distanz und dem Alter beim Erleben in Bezug bringen.

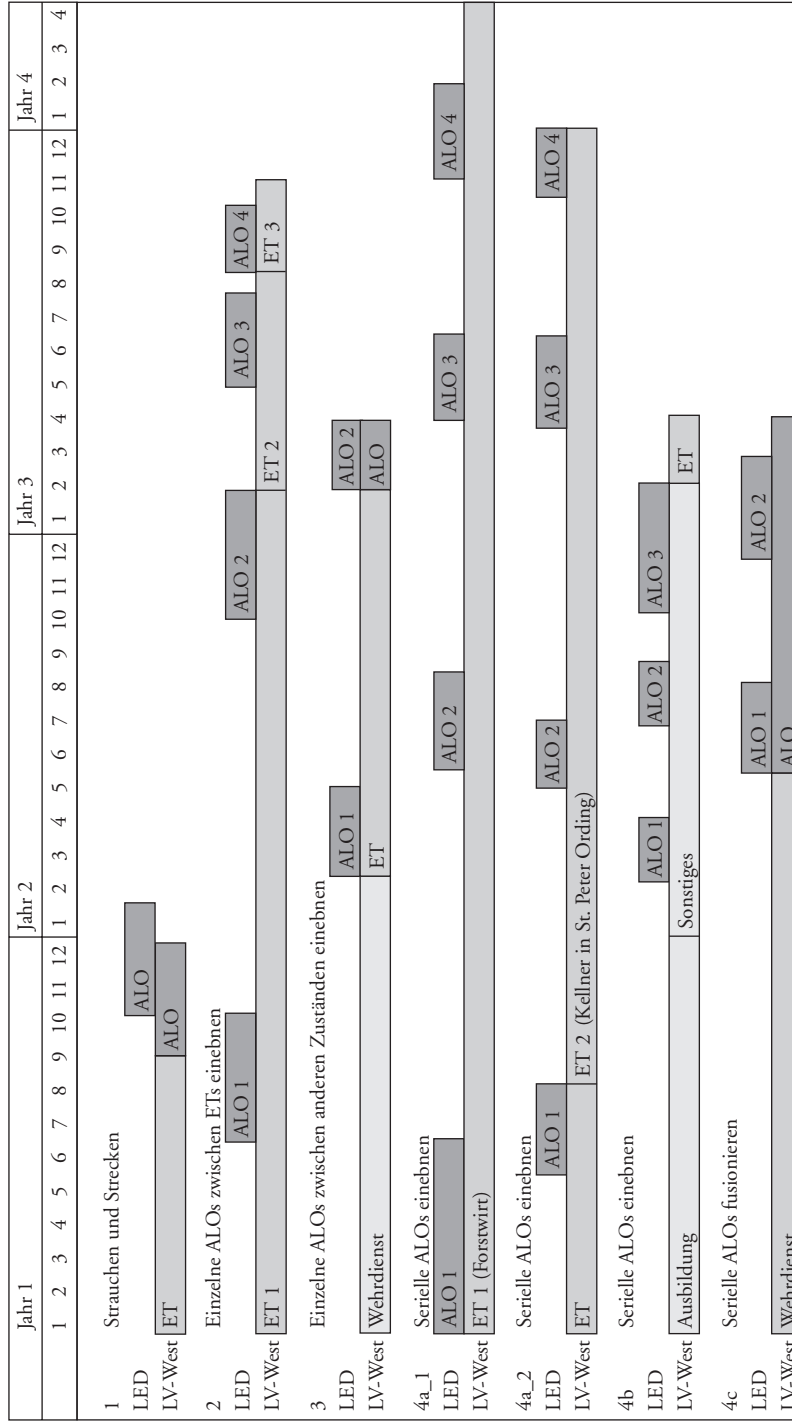
### 6.3 Auswertung I: Qualitative Einzelfallbetrachtung

An einer Stichprobe von 50 Personen, die in den beiden Datensätzen deutlich unterschiedlich viel Zeit in Arbeitslosigkeit verbrachten, habe ich sieben Arten der Abweichung identifiziert, die aus den unterschiedlichen Datenkonstruktionsprozessen und unterschiedlichen Konzepten von Arbeitslosigkeit und Leistungsbezug erklärt werden können. Diese sind in Abbildung 6.4 anhand ausgewählter Einzelfälle veranschaulicht<sup>7</sup>. Teilweise entsprechen sie den Abweichungen, die im Kapitel 5 identifiziert wurden. Da der Vergleich vollständige Sequenzen betrifft, die von der ersten Arbeitslosigkeit nach Abschluss der allgemein bildenden Schulen bis zur Gegenwart erfragt wurden, entstehen hier keine Abweichungen durch das Hinein- und Hinausdatieren von Episoden in bzw. aus einem begrenzten Zeitfenster.

<sup>7</sup> Freilich verbleiben Abweichungen, die keinem der Phänomene eindeutig zugeordnet werden können, exemplarisch dargestellt in Fall 8. Ich gehe aber davon aus, dass sich auch diese Fälle prinzipiell bei besserem Verständnis der individuellen und kontextuellen Hintergründe einem der Phänomene zuordnen ließen.



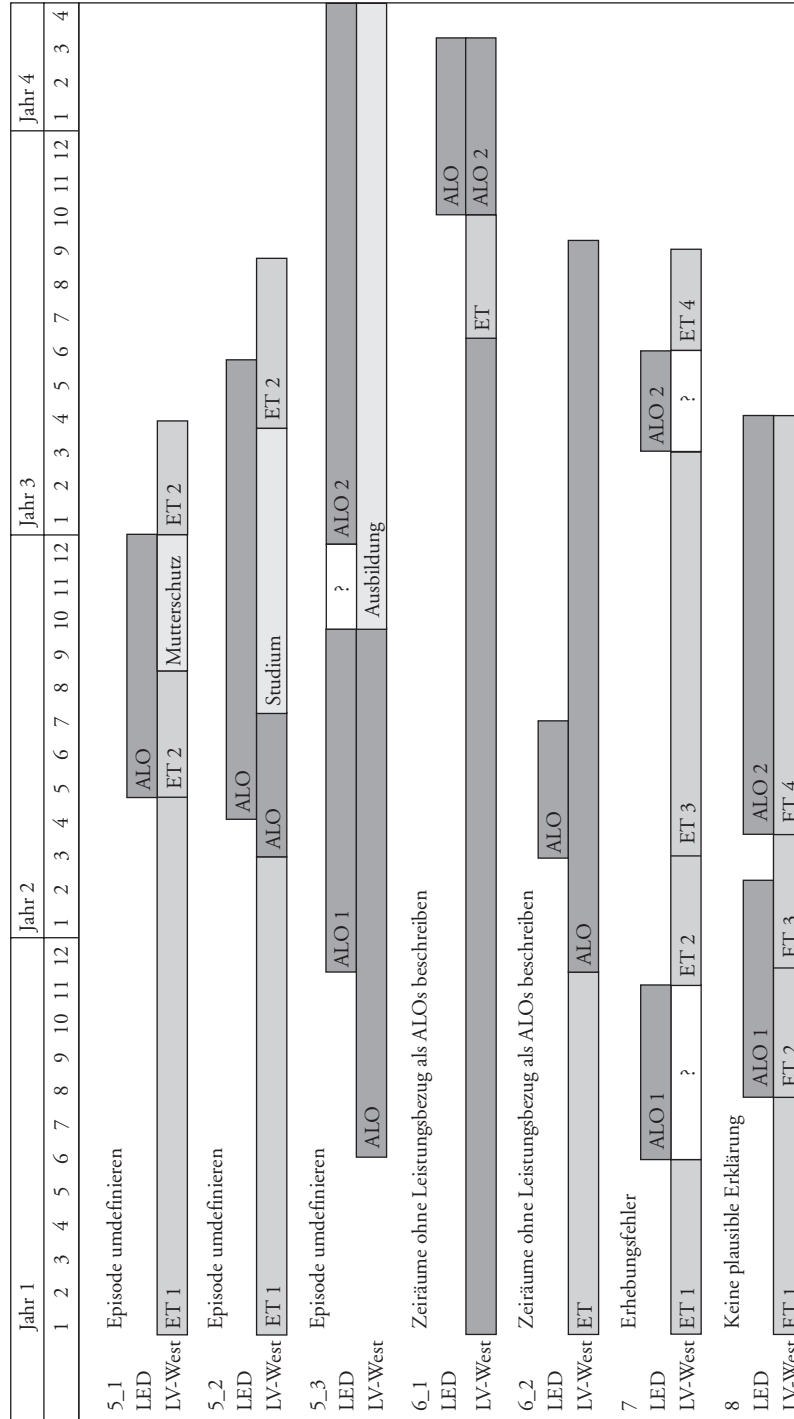
Abbildung 6.4: Sieben Arten der Abweichung zwischen LV-West und LED



ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

noch Abbildung 6.4: Sieben Arten der Abweichung zwischen LV-West und LED



ET = Erwerbstätigkeit; ALO = Arbeitslosigkeit.  
 Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

*Fall 1: Strecken/Stauchen/Verschieben*

Der Befragte rekonstruiert einen Lückenzeitraum, den er als Arbeitslosigkeit beschreibt und der sich mehr oder weniger mit einem Leistungsbezugszeitraum in der LED deckt. Lediglich der Beginn der Episode ist einen Monat früher und das Ende zwei Monate später datiert. Diese Abweichung beruht also auf einem Fehler auf der Stufe drei des Rekonstruktionsprozesses. Auch durch eine Sperrzeit des Arbeitsamts, das bei eigenverantwortlicher Kündigung des Betroffenen erst nach mehreren Wochen Zahlungen erteilt, oder durch Rundungsunschärfen beim Anpassen der Datumsangaben der LED können solche Abweichungen entstehen.

*Fall 2: Einzelne Arbeitslosigkeiten zwischen Erwerbstätigkeiten einebnen*

Der Befragte unterschlägt in drei Jahren vier Leistungsbezugszeiträume von zwischen 1 und 201 Tagen Dauer und füllt die entsprechenden Zeiträume durch das Strecken angrenzender Erwerbstätigkeiten. Die erste und dritte Leistungsbezugsepisode werden spurlos eingeebnet; mit den Enden der unterschlagenen Arbeitslosigkeitsepisoden 2 und 4 hingegen fallen zwei der Übergänge zwischen den berichteten Erwerbstätigkeiten zusammen – der Befragte rekonstruiert also hier in Übereinstimmung mit den Daten in der LED eine Veränderung in zentralen Aktivitäten, sozialen Rollen, beteiligten Personen und Orten. Diese Veränderungen werden aber als Veränderungen zwischen Erwerbstätigkeiten rekonstruiert; die Leistungsbezugszeiträume dazwischen sind keine Lebensabschnitte in der Repräsentationsstruktur der Erwerbsgeschichte.

*Fall 3: Einzelne Arbeitslosigkeiten zwischen anderen Zuständen einebnen*

Der Befragte unterschlägt einen Monat Leistungsbezug, der sich genau an dem von ihm berichteten Übergang vom Wehrdienst in die erste Erwerbstätigkeit zutrug. Der erinnerte Wechsel in den sozialen Rollen und Aktivitäten ist der zwischen Wehrdienst und Erwerbstätigkeit; die kurze Arbeitslosigkeit stellte also keinen Bruch dar und ist nicht als Lebensabschnitt repräsentiert. Vergleichbar tritt dieses Phänomen bei anderen Personen zwischen Ausbildungs- und Erwerbsepisoden auf, teilweise im Anschluss an die Erstausbildung und teilweise später im Anschluss an weitere Ausbildungen.

*Fall 4: Serielle Arbeitslosigkeiten einebnen*

Wo Leistungsbezugszeiträume in Serie auftreten, besteht ein erhöhtes Risiko, dass sie eingeebnet werden, weil sie oft kurz sind und weil der bruchhafte Charakter der einzelnen Episode durch ihr gewohnheitsmäßiges Auftreten ebenfalls reduziert wird. Wenn sie repräsentiert werden, werden sie aufgrund ihrer Ähnlichkeit und zeitlichen Nähe als generische Repräsentation gespeichert, sodass einzelne Episoden und Übergänge nur noch schwer rekonstruierbar und datierbar sind. Drei Phänomene treten deshalb bei der Repräsentation von Zeiträumen mit seriellem Leistungsbezug auf: das Einebnen durch Erwerbstätigkeiten (4a), das Einebnen durch andere Aktivitäten (4b) und das Fusionieren von Arbeitslosigkeiten und Einebnen der dazwischenliegenden Zeiträume (4c).

Die Personen 4a\_1 und 4a\_2 haben über Jahre hinweg jeden Winter für einen bis maximal drei Monate Leistungen bezogen. Die Berufe der Betroffenen (Forstwirt bzw. Kellner und Restaurantleiter in einem Ausflugsort an der Nordsee) legen nahe, dass es sich hierbei um

saisonal bedingte Entlassungen handelte und über die Fortsetzung der Erwerbsarbeit im Anschluss wenig Zweifel bestand. Insofern fehlen diesen Zeiträumen die Voraussetzungen für die Repräsentation als Lebensabschnitte in Arbeitslosigkeit.

Person 4b beschreibt einen Zeitraum zwischen Ausbildung und Erwerbstätigkeit, in dem sie innerhalb von etwa anderthalb Jahren dreimal für einen bis drei Monate Leistungen bezog, als „sonstiges“. Innerhalb dieser Zeit werden also erstens keinerlei nennenswerte Veränderungen in den zentralen Rollen und Aktivitäten repräsentiert oder rekonstruiert, und zweitens wird für den aktivitäts- und rollenhomogenen Zeitraum Arbeitslosigkeit nicht als die angemessenste Kategorie erachtet, obwohl Leistungsbezugszeiträume knapp die Hälfte ausmachen.

Bei Person 4c hingegen, die ebenfalls etwa für die Hälfte des Zeitraums Leistungen bezieht, „obsiegt“ die Arbeitslosigkeit: Zwei serielle Arbeitslosigkeiten werden fusioniert und der zwischenzeitliche Übergang in einen anderen Zustand und wieder zurück in Arbeitslosigkeit nicht rekonstruiert; der Befragte erinnert sich an keine Veränderungen in den sozialen Rollen und Aktivitäten innerhalb dieses Zeitraums. Da in der LED keine Angaben zu anderen Zuständen vorliegen, ist leider nicht klärbar, was in diesem Zeitraum vor sich ging; es mag eine geringfügige Beschäftigung gewesen sein, eine Leistungssperre (allerdings ist der Zeitraum hier mit über vier Monaten länger als die maximale Sperrzeit von zwölf Wochen), eine Aus- oder Weiterbildung oder eine längere Krankheit oder Rehabilitationsmaßnahme.

#### *Fall 5: Episode umdefinieren*

Wenn während eines Leistungsbezugszeitraums Rollen und Aktivitäten im Vordergrund stehen, spielt die Tatsache des Leistungsbezugs keine Rolle dafür, wie der entsprechende Zeitraum repräsentiert und rekonstruiert wird. Im Fall 5\_1 hat die Befragte übereinstimmend mit der LED rekonstruiert, dass und wann Veränderungen in zentralen erwerbs- und arbeitsmarktbezogenen Rollen, Aktivitäten usw. stattfanden, sie beschreibt den Leistungsbezugszeitraum aber als Erwerbstätigkeit (als stundenweise Aushilfe in einem Café mit wechselnden Arbeitszeiten und einem Monatseinkommen von 200 DM) und die letzten beiden Monate als Mutterschutz.

Fall 5\_2 war offenbar während des Leistungsbezugszeitraums ein Semester lang an einer Hochschule eingeschrieben; wenn er in diesem Zeitraum tatsächlich Veranstaltungen besucht und unter Umständen sogar das Ziel eines Studienabschlusses verfolgt hat, ist leicht verständlich, dass diese Zeit nach ihren Aktivitäten und sozialen Rollen als Lebensabschnitt in Ausbildung rekonstruiert wurde. Fall 5\_3 hat ebenfalls eine Vollzeitausbildung für einen Zeitraum des Leistungsbezugs angegeben; wenn es sich hier nicht um unrechtmäßigen Leistungsbezug handelt, ist anzunehmen, dass es sich bei der Leistung, deren Bezug hier im LED vermerkt ist, nicht um Arbeitslosengeld oder -hilfe, sondern um Unterhaltsgeld handelt.

#### *Fall 6: Zeiträume ohne Leistungsbezug als Arbeitslosigkeit beschreiben*

Die Erhebung von Arbeitslosigkeit ohne nähere Definition des Begriffs führt auch dazu, dass Zeiträume, die in der LV-West als Arbeitslosigkeit beschrieben werden, im LED keinerlei (Fall 6\_1) oder nur teilweise (Fall 6\_2) Entsprechung finden. Da in der LED keinerlei Information zu anderen Zeiträumen vorliegt, ist leider nicht zu klären, welchen Aktivitäten diese Personen hier im Einzelnen nachgegangen sind und ob dieser Zeitraum wirklich am angemessensten als Arbeitslosigkeit beschrieben wird.

*Fall 7: Erhebungsfehler*

Der Befragte hat zweimal am Ende einer Erwerbsepisode eine vorübergehende Unterbrechung seiner Erwerbstätigkeit angegeben und diesen Übergang in eine Lücke auch genau dort datiert, wo auch der Leistungsbezugszeitraum beginnt. Dann aber wurde der Befragte vermutlich aufgrund eines Steuerungsfehlers der CATI-Prozedur zu diesem Zeitraum nicht näher befragt, sodass diese Zeiträume keinem der Zustände zugewiesen werden können. In der Edition werden solche Lücken durch intensive Nachrecherche geschlossen. Da es sich bei den verwendeten LV-Daten aber um die unedierte Version handelt, finden sich vereinzelt Zeiträume, für die keinerlei Angaben vorliegen. Solche Abweichungen sind nicht kognitiv verursacht und stehen vermutlich auch nicht in systematischem Zusammenhang mit Unterschieden in den Datengenerierungsprozessen.

Die Einzelfallanalyse hat gezeigt, dass durch die dreistufige retrospektive Konstruktion nicht alle Leistungsbezugszeiträume in der Befragung als Arbeitslosigkeiten rekonstruiert werden und dass nicht alle in der Befragung rekonstruierten Arbeitslosigkeiten mit Leistungsbezugszeiträumen einhergehen. Dabei spielt die dritte Stufe des Datengenerierungsprozesses, das Datieren, die geringste Rolle, da durch die Betrachtung der Gesamtsequenz Episoden nicht durch Datierungsfehler ganz aus dem Betrachtungszeitraum herausfallen können. Im nächsten Abschnitt werde ich nun die Abweichungen quantitativ beschreiben und mit der Länge der retrospektiven Distanz in Verbindung bringen.

**6.4 Auswertung II: Konsistenz auf Monats- und Episodenebene****6.4.1 Vorgehen und Überblick über die Daten**

Um die tagesgenauen Angaben in der LED mit den monatsgenau erfassten Arbeitslosigkeitsphasen in der LV-West vergleichbar zu machen, werden nur Leistungsbezugszeiträume ab einem halben Monat, also 16 Tagen berücksichtigt und kürzere Episoden gelöscht. Dann behandle ich jeden Monat, der mindestens 16 Tage Leistungsbezug enthält, als Monat mit Leistungsbezug – unabhängig davon, auf wie viele Episoden diese Tage verteilt sind.

Tabelle 6.1 gibt einen Überblick über das Datenmaterial in der LED und der LV-West. Die Leistungsbezugszeiträume umfassen etwa 350 Monate mehr, als von den in der Befragung berichteten Arbeitslosigkeitsepisoden abgedeckt werden. In der LED formieren sich diese zu mehr als doppelt so vielen Episoden wie in der LV-West, die im Durchschnitt deutlich kürzer sind. Und auch die gesamten Erwerbsbiografien stellen sich in den beiden Datensätzen unterschiedlich dar: Besonders auffällig ist der in der LV-West deutlich höhere Anteil an Personen ohne Arbeitslosigkeit und der deutlich geringere Anteil an Personen, die ein bis sechs Monate in Arbeitslosigkeit verbrachten. Dieser erste Vergleich suggeriert, dass in der LV-West kurze Leistungsbezugszeiträume nicht als Arbeitslosigkeit rekonstruiert und repräsentiert werden, was zur Unterrepräsentation von Arbeitslosigkeit in den Befragungsdaten führt.

Da Vergleiche auf der Stichprobenebene nur begrenzt aussagekräftig sind, werde ich ausgehend vom Monat als kleinster, beiden Datensätzen gemeinsamen, Datenbezugseinheit Vergleiche auf der Individualebene durchführen. Jeder Monat innerhalb der Erwerbsgeschichte einer Person kann in der LED mit Leistungsbezug einhergehen oder nicht und in der LV-

Tabelle 6.1: Arbeitslosigkeitsepisoden in der LV-West und der LED

	LED		LV-West	
	absolut	in %	absolut	in %
Monate	2.752		2.399	
<i>Episoden</i>				
<i>N</i>	577		252	
<i>Episodendauer</i>				
Durchschnitt (Monate)	4,77		9,52	
Range (Monate)	1–37		1–56	
<i>N Episoden mit Dauer</i>				
1 bis 3 Monate	341	59,1	43	17,1
3 bis 12 Monate	197	34,1	152	60,3
über 12 Monate	39	6,8	57	22,6
<i>Personen/Lebensläufe</i>				
<i>N Episoden – Spannbreite</i>	1–17		1–6	
<i>Kumulative Arbeitslosigkeitsdauer</i>				
0 Monate in ALO	381	59,9	459	72,2
1 bis 6 Monate in ALO	125	19,7	53	8,3
6 bis 12 Monate in ALO	57	9,0	59	9,3
12 bis 24 Monate in ALO	42	6,6	42	6,6
über 24 Monate in ALO	31	4,9	23	3,6

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

West als arbeitslos angegeben werden oder nicht. Dadurch ergeben sich für jeden Monat einer jeden Person vier mögliche Kombinationen:

- in LED und in LV-West übereinstimmend *nicht* arbeitslos (konsistent),
- in LED arbeitslos, in LV-West nicht (Inkonsistenz vom Typ A),
- in LV-West arbeitslos, aber in LED nicht (Inkonsistenz vom Typ B),
- in LED und in LV-West übereinstimmend arbeitslos (konsistent).

Ich beschränke mich in den folgenden Analysen auf die 282 Personen (44,5% der Ausgangsstichprobe), die mindestens für einen Monat Leistungen bezogen und/oder Arbeitslosigkeit berichtet haben, und schließe diejenigen aus, die im gesamten Zeitraum weder Leistungen bezogen noch Arbeitslosigkeiten berichten. Die Übereinstimmung hierbei ist in beide Richtungen relativ hoch: Von den 381 Personen ohne Leistungsbezug haben 92,9 Prozent auch keine Arbeitslosigkeit angegeben, und umgekehrt haben von den 459 Personen ohne Arbeitslosigkeitsangaben 77,1 Prozent auch keinerlei Leistungen bezogen.

Ein *Monat* kann aus mehreren Gründen inkonsistent sein: Entweder gehört er zu einer Episode, die im jeweils anderen Datensatz gänzlich unterschlagen wurde (z.B. Einzelfälle 2 und 4a), oder er gehört zu einer Episode, die nur teilweise mit dem jeweils anderen Datensatz übereinstimmt, weil sie bei der Rekonstruktion verschoben, gestreckt oder gestaucht wurde (Einzelfall 1), weil sie mit anderen Leistungsbezugszeiträumen zu einer größeren Episode fusioniert wurde (Einzelfall 4c), weil der Leistungsbezugszeitraum in der Rückschau nur zum Teil als Arbeitslosigkeit rekonstruiert wurde (Einzelfall 5\_2) oder weil die an den Zeitraum des Leistungsbezugs angrenzenden Monate ebenfalls als Arbeitslosigkeit rekonstruiert werden (Einzelfall 6\_2).

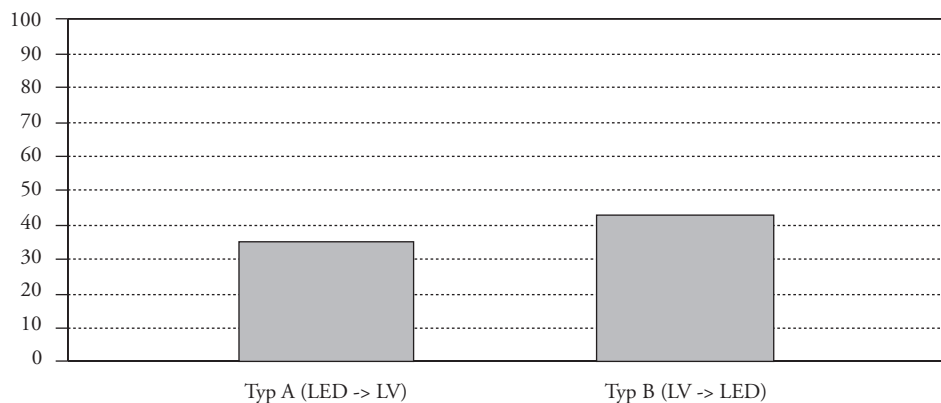
Da Episoden sich aus konsistenten und/oder inkonsistenten Monaten zusammensetzen, können sie hinsichtlich des *Grades* ihrer Übereinstimmung aus *wechselnder Perspektive* beschrieben werden. Zum einen kann angegeben werden, ob ein Leistungsbezugszeitraum gar nicht, teilweise oder in voller Länge einer Arbeitslosigkeit im Befragungsdatensatz entspricht (Inkonsistenz vom Typ A). Zum anderen kann angegeben werden, ob eine rekonstruierte Arbeitslosigkeit gar nicht, teilweise oder ganz mit Leistungsbezug einhergehen (Inkonsistenz vom Typ B). Für den Einzelfall 6\_2 kann angegeben werden, dass die drei Monate Leistungsbezug gänzlich auch als Arbeitslosigkeit rekonstruiert werden, dass aber umgekehrt die Arbeitslosigkeitsepisode von 20 Monaten nur teilweise mit Leistungsbezug einhergeht.

#### 6.4.2 Konsistenz auf Monats- und Episodenebene

In Abbildung 6.5 sind die Inkonsistenzen von Typ A und B beim Vergleich von Monaten dargestellt. Von 2.752 Monaten, für die in der LED Leistungsbezüge verzeichnet sind, wurde für etwa 43 Prozent auch Arbeitslosigkeit angegeben. Umgekehrt finden sich für fast genau die Hälfte der 2.399 Monate in Arbeitslosigkeit auch Leistungsbezüge in der LED.

In Abbildung 6.6 sind die Inkonsistenzen von Typ A und B auf Episodenebene dargestellt<sup>8</sup>. Über 60 Prozent der Leistungsbezugsepisoden enthalten keinen einzigen als arbeitslos angegebenen Monat. Da das Fehldatieren und Verschieben von Episoden nur eine gering-

Abbildung 6.5: Übereinstimmung zwischen LV-West und LED auf Monatsebene (in %)



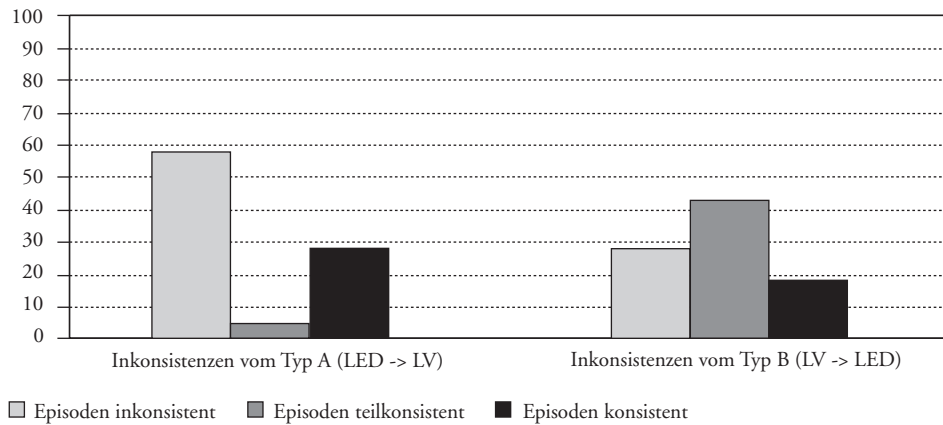
Typ A: Anteil der Monate mit Leistungsbezug, die auch in der LV-West als arbeitslos angegeben werden ( $n = 2.752$  Monate).

Typ B: Anteil der in der LV als arbeitslos angegebenen Monate mit Leistungsbezug ( $n = 2.399$  Monate).

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

<sup>8</sup> Als gänzlich konsistent wurden außerdem diejenigen Episoden klassifiziert, die länger als einen Monat sind und nur einen einzigen Monat enthielten, der im jeweils anderen Datensatz inkonsistent war, um den Rundungsunschärfen bei der Umrechnung der Datumsangaben der LED Rechnung zu tragen.

Abbildung 6.6: Übereinstimmung zwischen LV-West und LED auf Episodenebene (in %)



Typ A: Anteil der Leistungsbezugsepisoden, die in der LV-WEST gar nicht angegeben werden, teilweise übereinstimmend oder ganz übereinstimmend angegeben werden ( $n = 577$  Episoden).

Typ B: Anteil der Arbeitslosigkeiten in der LV, die in der LED gar nicht verzeichnet sind, teilweise übereinstimmend oder ganz übereinstimmend verzeichnet sind ( $n = 252$  Episoden).

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

für eine wichtige Rolle spielt, kann man davon ausgehen, dass diese Episoden gänzlich unterschlagen oder vergessen wurden. Nur ein geringer Anteil der Leistungsbezugsepisoden enthält konsistente und inkonsistente Monate, und 32,1 Prozent enthalten nur konsistente Monate.

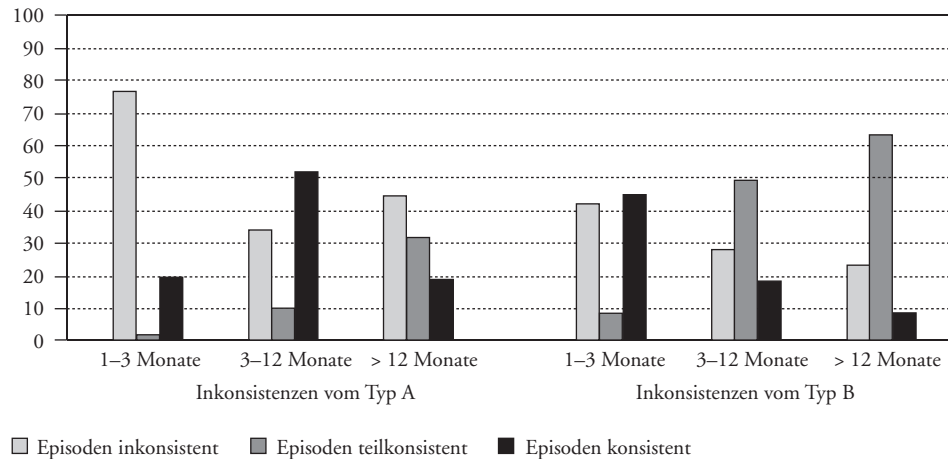
Umgekehrt enthält nahezu ein Drittel der in der Befragung angegebenen Arbeitslosigkeitsepisoden ausschließlich inkonsistente Monate. Da Verschieben und Fehldatieren nur eine geringe Rolle spielen, können wir davon ausgehen, dass es sich hierbei um hinzugefügte Episoden handelt. Über 40 Prozent der Episoden hingegen enthalten sowohl konsistente als auch inkonsistente Monate, und nur etwa ein Fünftel setzt sich nur aus konsistenten Monaten zusammen. Leistungsbezugszeiträume werden also fast immer entweder völlig unterschlagen oder zur Gänze konsistent als Arbeitslosigkeitsphasen rekonstruiert. Leistungsbezugsepisoden sind auch oft nur sehr kurz und können deshalb seltener in diese Kategorie fallen. Umgekehrt beinhalten rekonstruierte Arbeitslosigkeitsepisoden oft sowohl Leistungsbezugszeiträume als auch Zeiträume ohne Leistungsbezug.

### 6.4.3 Konsistenz und Episodenlänge

In Abbildung 6.7 ist der Zusammenhang von Episodenlänge und Grad der Übereinstimmung dargestellt, indem die Episoden in drei Dauerkategorien eingeteilt wurden: ein bis drei Monate, vier bis zwölf Monate oder länger als zwölf Monate. Was Inkonsistenzen vom Typ A betrifft, so werden die Leistungsbezugsepisoden von vier bis zwölf Monaten Dauer am ehesten in der LV-West rekonstruiert: Der Anteil an übereinstimmenden Episoden ist



Abbildung 6.7: Abhängigkeit der Übereinstimmung zwischen LV-West und LED von der Episodenlänge (in %)



Typ A: Anteil der Leistungsbezugsepisoden, die in der LV-West gar nicht, teilweise übereinstimmend oder ganz übereinstimmend angegeben werden ( $n = 577$  Episoden).

Typ B: Anteil der Arbeitslosigkeiten in der LV, die in der LED gar nicht, teilweise übereinstimmend oder ganz übereinstimmend verzeichnet sind ( $n = 252$  Episoden).

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

höher und der Anteil an zu gänzlich unterschlagenen Episoden geringer als bei kürzeren und längeren Leistungsbezugszeiträumen. Bei Inkonsistenzen vom Typ B hingegen ist sowohl der Anteil der hinzugefügten Arbeitslosigkeiten als auch der Anteil der ganz konsistenten Episoden für die kürzesten Episoden am höchsten und beides sinkt mit zunehmender Episodenlänge. Hingegen steigt das Risiko für Teilkonsistenz. Das günstigste *Verhältnis* von hinzugefügten und konsistenten Arbeitslosigkeiten besteht allerdings für die kürzesten Episoden.

Sehr kurze Leistungsbezugszeiträume werden also in der Tat am häufigsten (in über drei Viertel der Fälle) gänzlich vergessen, und sehr kurze Arbeitslosigkeitsphasen finden ebenfalls am häufigsten keinerlei Entsprechung in der LED. Darüber hinaus aber ist das Verhältnis von Episodenlänge und Übereinstimmung aber komplexer als eingangs vermutet, besonders für Inkonsistenzen vom Typ B.

## 6.5 Auswertung III: Die retrospektive Distanz

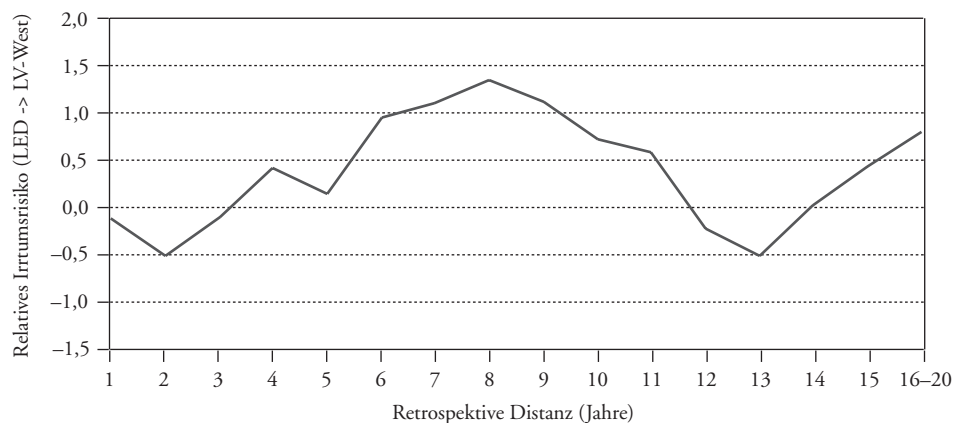
### 6.5.1 Das Unterschlagen von Leistungsbezugszeiträumen und die retrospektive Distanz

Um die Inkonsistenz vom Typ A für Monate mit der Länge der retrospektiven Distanz und dem Alter beim Erleben in Bezug zu setzen, betrachte ich für jedes verstrichene Jahr das

relative Risiko (*odds*) für einen Monat mit Leistungsbezug, abweichend davon in der LV-West nicht als Arbeitslosigkeit berichtet zu werden. Dieses Maß wird errechnet, indem die Wahrscheinlichkeit für das Auftreten des inkonsistenten Falls durch die Wahrscheinlichkeit für das Auftreten geteilt und der natürliche Logarithmus dieses Quotienten berechnet wird. Es gibt die Wahrscheinlichkeit an, dass eines von zwei Ereignissen (Monat inkonsistent) eintritt im Vergleich zu der Wahrscheinlichkeit, dass das andere Ereignis (Monat konsistent) eintritt. Positive Werte zeigen an, dass inkonsistente Paare häufiger sind als konsistente, und negative Werte zeigen an, dass konsistente Monate die inkonsistenten überwiegen.

Im Zeitraum von ein bis zwei Jahren vor der Befragung beispielsweise waren von allen Befragten 276 Monate mit Leistungsbezug in der LED verzeichnet. Davon waren 157 auch in der LV-West als arbeitslos angegeben worden und 89 nicht. Das relative Risiko für inkonsistente Angaben beträgt also  $89/157 = 0,57$ , was ein *log odds* von  $-0,56$  ergibt. Diese Zahl sagt zum einen aus, dass inkonsistente Monate in diesem Jahr seltener sind als konsistente, und zum anderen, dass das Irrtumsrisiko in diesem Jahr gegenüber der Chance auf Übereinstimmung um etwas mehr als die Hälfte niedriger ist<sup>9</sup>.

Abbildung 6.8: Einfluss der retrospektiven Distanz auf die Inkonsistenzen vom Typ A (LED -> LV-West) auf Monatsebene



Relatives Risiko (*log odds*) eines Monats mit Leistungsbezug, in der LV-West nicht arbeitslos zu sein. Positive Werte: mehr inkonsistente Monate, negative Werte: mehr konsistente Monate.  $n = 2.752$  Monate.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

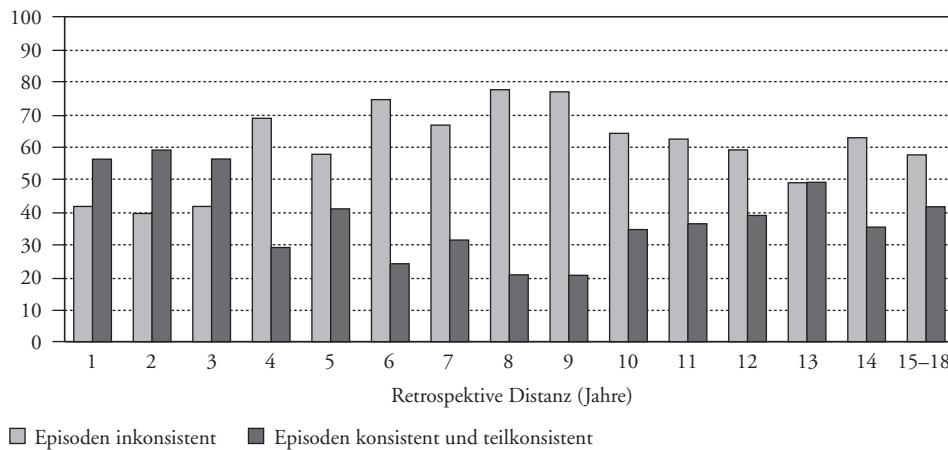
<sup>9</sup> Die Alternative wäre, Wahrscheinlichkeiten anzugeben, die sich als Anteil inkonsistenter Angabenpaare an allen Angabenpaaren berechnen; im angeführten Beispiel wäre dies eine Wahrscheinlichkeit von  $89/276 = 33,3$  Prozent. Der Vorteil der relativen Risiken ist, dass in ihnen Information über die Wahrscheinlichkeit eines Ereignisses *relativ zu einem anderen* enthalten ist, was im vorliegenden Fall besonders angemessen erscheint.

In Abbildung 6.8 sehen wir, wie sich das relative Risiko über die Jahre verändert. In den ersten drei Jahren ist ein u-förmiger Verlauf erkennbar: Im Jahr vor der Befragung kommen konsistente und inkonsistente Monate etwa gleich häufig vor, im zweiten sind konsistente Monate deutlich häufiger und ab dem dritten Jahr überwiegen die inkonsistenten. Danach steigt das Irrtumsrisiko weiter bis acht Jahre retrospektiver Distanz, wo auf 35 konsistente 129 inkonsistente Monate kommen. Danach sinkt es allerdings wieder, bis bei etwa 12 Jahren retrospektiver Distanz konsistente und inkonsistente wieder etwa gleich häufig und bei 13 Jahren konsistent sogar wieder häufiger vorkommen. Erst ab 15 Jahre steigt das Irrtumsrisiko dann wieder an.

Ob die Inkonsistenz vom Typ A auf Monatesebene über die Zeit zunimmt, werde ich betrachten, indem ich die Verteilung von konsistenten, teilkonsistenten und inkonsistenten Episoden innerhalb eines jeden Jahres vergleiche. Da teilkonsistente Episoden in dieser Vergleichsrichtung kaum vorkommen, wurden teilkonsistente und konsistente Episoden zusammengefasst. Da Episoden eine zeitliche Ausdehnung haben, wurde jede Episode dem Jahr zugeschlagen, in dem ihr zeitlicher Mittelpunkt liegt; eine Episode, die von November 1990 bis Mai 1991 reicht, wird also dem Jahr 1990 zugeordnet.

Abbildung 6.9 zeigt ein ähnliches Bild wie beim Vergleich auf Monatesebene: Der Anteil gänzlich unterschlagener Leistungsbezugsepisoden steigt nach den ersten drei Jahren, in denen er bei etwa 40 Prozent liegt, mehr oder wenig kontinuierlich an. Komplementär dazu sinkt der Anteil konsistenter und teilkonsistenter Episoden und erreicht seinen Tiefstand bei etwa acht Jahren vor der Befragung. Bei 13 Jahren Rückschau liegen konsistente und inkonsistente Episoden wieder gleichauf, kurzzeitig überwiegen noch einmal die konsistenten Episo-

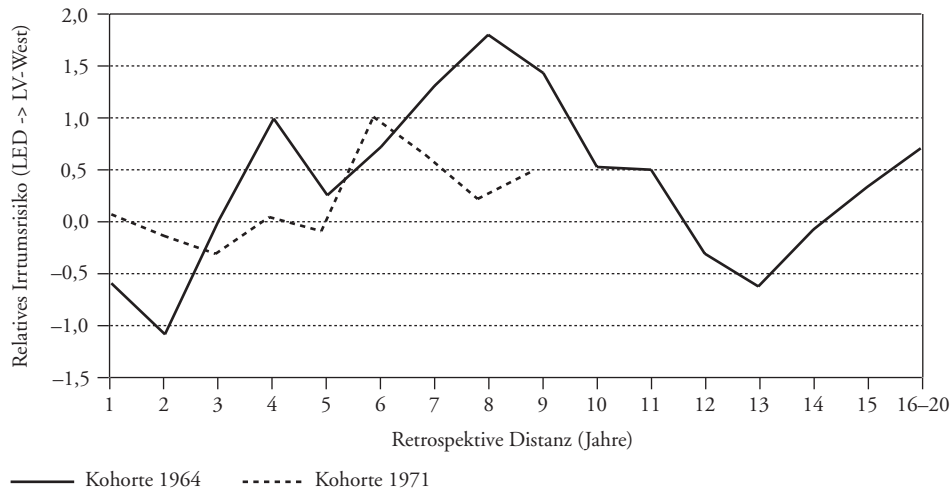
Abbildung 6.9: Einfluss der retrospektiven Distanz auf die Inkonsistenzen vom Typ A (LED -> LV-West) auf Episodenebene (in %)



Typ A: Anteil der Leistungsbezugsepisoden, die in der LV-WEST gar nicht angegeben werden, teilweise übereinstimmend oder ganz übereinstimmend angegeben werden ( $n = 577$  Episoden).

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

Abbildung 6.10: Einfluss der retrospektiven Distanz auf die Inkonsistenzen vom Typ A (LED -> LV-West) auf Monatsebene getrennt nach Kohorten



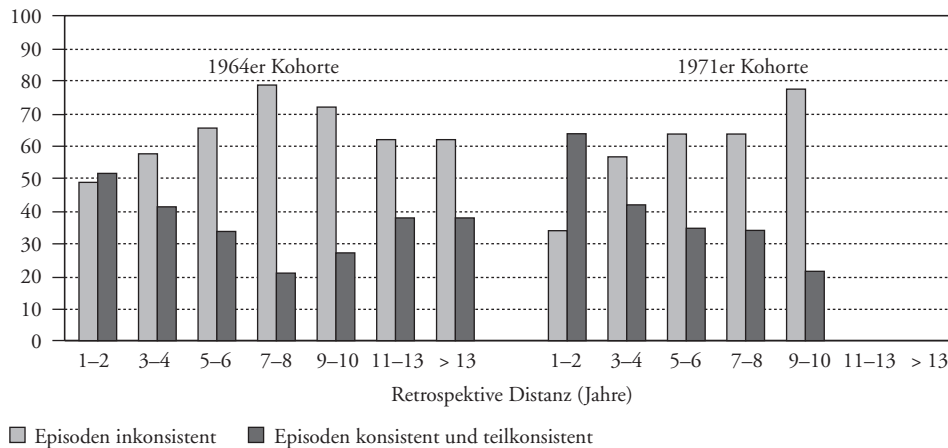
Relatives Risiko (*log odds*) eines Monats mit Leistungsbezug, in der LV-West nicht arbeitslos zu sein. Positive Werte: mehr inkonsistente Monate, negative Werte: mehr konsistente Monate.  $n = 2.752$  Monate (Kohorte 1964:  $n = 1.766$ , Kohorte 1971:  $n = 986$ ).

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

den, bis für die am weitesten zurückliegenden Jahre wieder deutlich die inkonsistenten Episoden überwiegen. Die Übereinstimmung verändert sich auch hier über den Vergleichszeitraum, nicht aber in der Form einer kontinuierlichen Verschlechterung.

Da in der LV-West zwei Geburtsjahrgänge – 1964 und 1971 – befragt wurden, fällt für die Befragten dieselbe retrospektive Distanz mit unterschiedlichen Lebensaltern zusammen; fünf Jahre vor der Befragung waren die Angehörigen der älteren Kohorte etwa 29 Jahre alt, die Angehörigen der jüngeren Kohorte hingegen etwa 22 Jahre (siehe Abb. 6.1). Dadurch habe ich die Möglichkeit, Einflüsse des Alters beim Erleben einer Arbeitslosigkeit von denen der retrospektiven Distanz zu trennen. In Abbildung 6.10 ist die Abhängigkeit des Irrtumsrisikos von der retrospektiven Distanz für die beiden Kohorten getrennt dargestellt. Läge ein Effekt des Alters beim Erleben (statt der retrospektiven Distanz) vor, müsste die Kurve der 1971er Kohorte der Kurve der 1964er Kohorte vom Verlauf her gleichen, aber um sieben bis acht Jahre versetzt verlaufen. Die Kurve der 1971er Kohorte verläuft zwar prinzipiell ähnlich wie die der 1964er Kohorte, steigt und fällt aber weniger extrem und fällt nach dem Gipfel früher wieder ab. Eine Verschiebung um sieben bis acht Jahre ist nicht auszumachen. Auch bei den Episoden (Abb. 6.11) verlaufen die Kurven der beiden Kohorten ähnlich; das Alter beim Erleben scheint also keine Rolle über die retrospektive Distanz hinaus zu spielen.

Abbildung 6.11: Einfluss der retrospektiven Distanz auf die Inkonsistenzen vom Typ A (LED -> LV-West) auf Episodenebene getrennt nach Kohorten (in %)



Typ A: Anteil der Leistungsbezugsepisoden, die in der LV-West gar nicht, teilweise übereinstimmend oder ganz übereinstimmend angegeben werden;  $N = 577$  Episoden (Kohorte 1964:  $n = 372$ , Kohorte 1971:  $n = 205$ ).

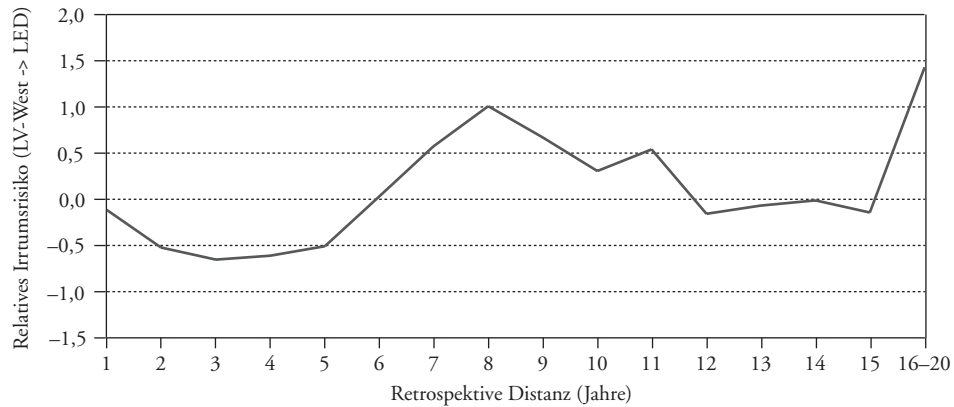
Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

### 6.5.2 Die Rekonstruktion von Arbeitslosigkeiten ohne Leistungsbezug und die retrospektive Distanz

In Abbildung 6.12 ist das relative Risiko eines in der LV-West arbeitslosen Monats, inkonsistent statt konsistent auszufallen, gegen die Zeit aufgetragen. Im Jahr vor der Befragung sind konsistente und inkonsistente Monate etwa gleich häufig; bis etwa sechs Jahre vor der Befragung überwiegen die konsistenten Monate, wobei auch hier ein u-förmiger Verlauf sichtbar wird. Danach kehrt sich das Verhältnis um, bis wiederum im Zeitraum bis etwa acht Jahre vor der Befragung auf 96 konsistente 39 inkonsistente Monate kommen. Danach sinkt das Irrtumsrisiko wieder auf ein ausgewogenes Verhältnis bei 12 bis 13 Jahren, um für die am weitesten zurückliegenden Jahre wieder steil anzusteigen.

In Abbildung 6.13 sind die Kurven getrennt nach Kohorten und nicht für die retrospektive Distanz, sondern für das (jahresweise) Alter beim Erleben aufgetragen. Da alle Befragten entweder im Jahre 1964 oder im Jahre 1971 geboren worden waren, sind retrospektive Distanz und Alter bei Erleben hoch korreliert, weichen aber um einen fast konstanten Betrag voneinander ab: Eine retrospektive Distanz von fünf Jahren bedeutet für einen Befragten aus der Kohorte 1964, dass er im betreffenden Jahr etwa 28 Jahre alt war, für einen Befragten aus der Kohorte 1971 hingegen, dass er etwa 23 Jahre alt war. Deshalb nehmen die Kurven einen ähnlichen Verlauf, aber um sieben bis acht Jahre versetzt. Bis auf ein drastisches Auseinanderklaffen im Alter von etwa 24 Jahren verlaufen die Kurven nahezu identisch. Dies spricht dafür, dass der Verlauf, insbesondere die massive Nichtübereinstimmung

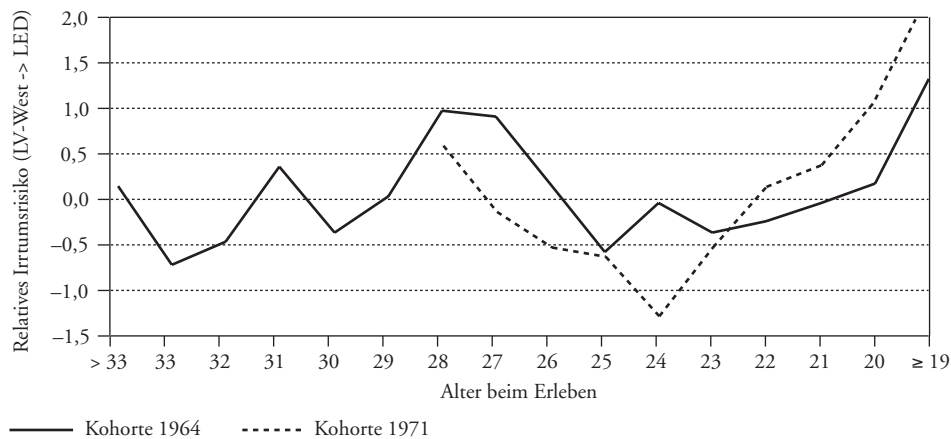
Abbildung 6.12: Einfluss der retrospektiven Distanz auf die Inkonsistenzen vom Typ B (LV-West -> LED) auf Monatebene



Typ B: Relatives Risiko (*log odds*) eines in der LV-West als arbeitslos angegebenen Monats, nicht mit Leistungsbezug in der LED verzeichnet zu sein ( $n = 2399$  Monate). Positive Werte: mehr inkonsistente Monate; negative Werte: mehr konsistente Monate.

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

Abbildung 6.13: Einfluss des Alters zum Zeitpunkt einer Arbeitslosigkeit auf die Inkonsistenzen vom Typ B (LV-West -> LED) auf Monatebene



Typ B: Relatives Risiko (*log odds*) eines in der LV-West als arbeitslos angegebenen Monats, nicht mit Leistungsbezug in der LED verzeichnet zu sein. Positive Werte: mehr inkonsistente Monate; negative Werte: mehr konsistente Monate.

$n = 2.399$  Monate (Kohorte 1964:  $n = 1.491$ , Kohorte 1971:  $n = 908$ ).

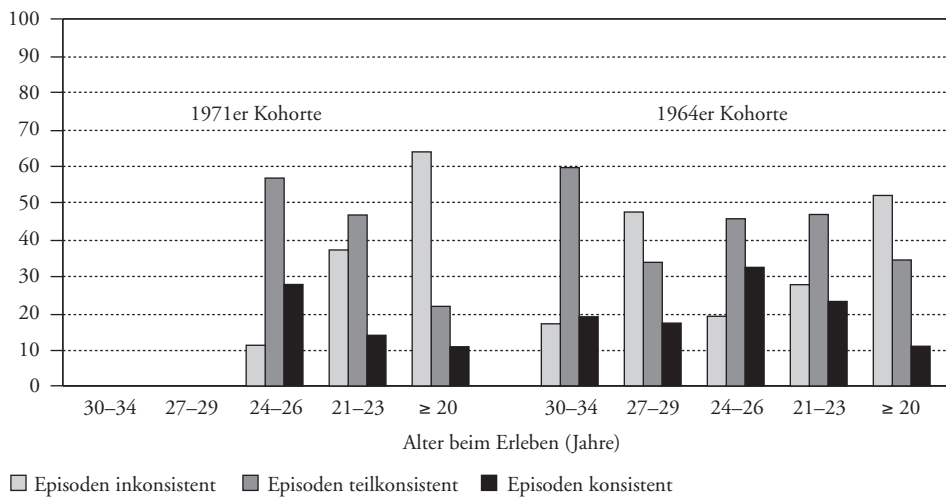
Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

am hinteren Ende des Zeitraums, eher mit dem Alter beim Erleben als mit der retrospektiven Distanz zusammenhängt.

In Abbildung 6.14 betrachte ich die Übereinstimmung auf Episodenbasis getrennt nach Kohorten und in Abhängigkeit vom Alter beim Erleben, wobei aufgrund der geringen Fallzahlen immer drei Jahre zusammengelegt werden. Hier machen teilkonsistente Episoden einen beträchtlichen Anteil von mindestens einem Drittel und teilweise über die Hälfte der Episoden innerhalb jeden Jahres aus. Ansonsten wiederholen sich die Ergebnisse des Monatsvergleichs: Innerhalb beider Kohorten ist bis 20 Jahre der Anteil von inkonsistenten Episoden am höchsten und sinkt, je älter die Befragten zum Zeitpunkt der jeweiligen Arbeitslosigkeit waren, während der Anteil an konsistenten und teilkonsistenten Episoden gegenläufig steigt. Ab dem Alter 27 bis 29 dreht sich dieser Trend um, und der Anteil der inkonsistenten Episoden sinkt wieder (für die 1971er Kohorte liegen hier keine Beobachtungen mehr vor), um allerdings für die Jahre über 30 wieder stark anzusteigen. Der Anteil konsistenter Episoden sinkt komplementär und bleibt bis zum Ende niedrig.

Inkonsistenzen vom Typ A und vom Typ B verändern sich über den Vergleichszeitraum; die Zu- bzw. Abnahme der inkonsistenten im Vergleich zu den konsistenten Angaben folgt dabei aber keiner kontinuierlichen Vergessenskurve: Weder das absolute noch das relative Risiko für Inkonsistenz nimmt über die etwa 15 Jahre stetig zu. Stattdessen zeigt sich ein diskontinuierlicher Verlauf, der für die Inkonsistenzen vom Typ B vom Alter beim Erleben mit gestaltet wird.

Abbildung 6.14: Einfluss des Alters zum Zeitpunkt einer Arbeitslosigkeit auf die Inkonsistenzen vom Typ B (LV-West -> LED) auf Episodenebene (in %)



Typ B: Anteil der Arbeitslosigkeiten in der LV-West, die in der LED gar nicht, teilweise übereinstimmend oder ganz übereinstimmend verzeichnet sind;  $N = 252$  Episoden (Kohorte 1964:  $n = 145$ , Kohorte 1971:  $n = 107$ ).

Quelle: Deutsche Lebensverlaufsstudie, Max-Planck-Institut für Bildungsforschung, Berlin, eigene Berechnungen; Leistungsempfängerdatei der Bundesanstalt für Arbeit, eigene Berechnungen.

## 6.6 Diskussion

### *Arbeitslosigkeit in der retrospektiven Erwerbsgeschichte*

Die Analysen haben wie erwartet gezeigt, dass Zeiträume des Leistungsbezugs und Arbeitslosigkeit in der retrospektiven Selbstdefinition nicht immer miteinander übereinstimmen. Die Abweichungen lassen sich durch die unterschiedlichen Datengenerierungsprozesse erklären, zu denen in den Befragungsdaten auch die Rekonstruktionsprozesse des autobiografischen Erinnerns gehören.

Leistungsbezugszeiträume umfassen eine Vielzahl sehr unterschiedlicher Episoden der Erwerbsabstinenz: saisonal bedingte Entlassungen (etwa im Bau- oder Gastronomiegewerbe) mit praktischer Wiedereinstellungsgarantie, unangenehme, aber nicht unerwartete Übergangszeiträume zwischen Ausbildung und Erwerbstätigkeit und schwerwiegende biografische Brüche und Verunsicherungen. Nicht alle davon sind auch in der Lebensgeschichte als Lebensabschnitte der Arbeitslosigkeit repräsentiert und werden ohne nähere Definition deshalb auch nicht verlässlich als solche rekonstruiert.

Die Repräsentation von Zeiträumen in der Lebensgeschichte als „arbeitslos“ wird nicht „ohne Not“ vorgenommen; ihre „Existenzberechtigung“ (bzw. Repräsentationsberechtigung) gewinnen Arbeitslosigkeiten vor allem als Brüche in der Erwerbsbiografie, als Wendepunkte für eine Neuorientierung des Lebens oder als überwundene Stolpersteine bei der erfolgreichen Rückkehr in die Erwerbsarbeit<sup>10</sup>. Sonst werden die betreffenden Leistungsbezugszeiträume eingebnet, oder wenn für die entsprechenden Zeiträume andere Aktivitäten und Rollen vorliegen, werden diese zur Beschreibung bevorzugt. Wie bereitwillig dies getan wird, illustriert der Einzelfall 5\_1, der für einen Zeitraum des Leistungsbezugs als Hauptaktivität lieber eine Erwerbstätigkeit als Aushilfe in der Gastronomie für 200 DM im Monat angab. Deshalb werden am ehesten Leistungsbezugszeiträume von mittlerer Länge (vier bis zwölf Monate) auch als Arbeitslosigkeit repräsentiert. Solche Leistungsbezugszeiträume sind lang genug, um prekär zu sein und einen Bruch in der Erwerbsbiografie zu verursachen, aber nicht so lang, dass in ihnen andere Aktivitäten als die Arbeitssuche im Vordergrund stehen oder sie umgedeutet und als Lebensabschnitt mit definitorischen Aktivitäten und Rollen repräsentiert oder rekonstruiert werden müssen. Außerdem stellen Leistungsbezugszeiträume von über einem Jahr Dauer einen Sonderfall dar: Eine derart lange Zahlung bedeutet, dass nach dem Auslaufen des Arbeitslosengeldes Arbeitslosenhilfe in Anspruch genommen oder dass Unterhaltsleistungen bezogen wurden.

Sich arbeitslos gemeldet zu haben, Leistungen zu beziehen und die vom Arbeitsamt geforderten Suchaktivitäten an den Tag zu legen und nachzuweisen, ist keine notwendige Voraussetzung dafür, dass Zeiträume in der ziel- und entscheidungsgetriebenen Entwicklungslogik als Lebensabschnitte in Arbeitslosigkeit repräsentiert zu werden. Deshalb kam es auch in beträchtlichem Umfang dazu, dass Zeiträume als Arbeitslosigkeit rekonstruiert

<sup>10</sup> Dies gilt für westliche stabile Systeme, in denen Arbeitslosigkeit als individuelles Schicksal in der persönlichen Entwicklungsgeschichte aus derselben heraus begründet werden muss. Mutz (1996) belegt in qualitativen biografischen Interviews, dass in Ostdeutschland die Erfahrung von Massenarbeitslosigkeit als kollektives und historisch bedingtes Schicksal andere Begründungsmuster entstehen lässt. Die Befunde des Kapitels 5 deuten ebenfalls in diese Richtung. Allerdings stellt Mutz fest, dass Arbeitslosigkeit auch in der DDR bzw. den fünf neuen Bundesländern als Abwesenheit von Erwerbsarbeit gedacht wird.



wurden, die nicht oder nur teilweise mit Leistungsbezug einhergingen. Für die Repräsentation eines Zeitraums als Arbeitslosigkeit ist also weniger ausschlaggebend, welche Aktivitäten ausgeführt wurden, sondern seine *Bewertung als Arbeitslosigkeit im Rahmen der Lebensgeschichte*.

Diese Selektivität der retrospektiven Rekonstruktion führt dazu, dass sich Arbeitslosigkeit im Lebensverlauf unterschiedlich darstellt, je nachdem welche Datenquelle zu Grunde gelegt wird. Aufgrund der LV-West würde man Inzidenz und Frequenz von Arbeitslosigkeit im Lebenslauf der Befragten deutlich unterschätzen. Die berichteten Episoden sind allerdings im Schnitt fast doppelt so lang wie in der LED, decken nur etwa 12 Prozent weniger Monate ab und ein Drittel Personen mehr verbringt lange Zeiträume (über ein Jahr) in Arbeitslosigkeit. So wird die *Verbreitung* von Arbeitslosigkeit unterschätzt, dafür aber ihre *Gewichtigkeit* und *Bedeutung* im Lebenslauf tendenziell überschätzt.

#### *Länge der retrospektiven Distanz*

Die Abweichungen zwischen LED und LV-West lassen sich zwar als Ergebnis retrospektiver Datenkonstruktion beschreiben; dies bedeutet jedoch nicht, dass sie mit der Länge der Rückschau kontinuierlich zunehmen. Die Veränderungen der Übereinstimmung über den Vergleichszeitraum lassen sich teilweise als Ergebnis des Einflusses mehrerer Größen erklären. Der u-förmige Verlauf bei der Abbildung der Leistungsbezugszeiträume in den Befragungsdaten (Inkonsistenzen vom Typ A) kann als Ausdruck einer Kombination aus sozialer Erwünschtheit und Vergessen interpretiert werden: Die Leistungsbezugszeiträume der Gegenwart und unmittelbaren Vergangenheit werden eher unterschlagen, weil die Befragten ungern zugeben, arbeitslos zu sein, und weiter zurückliegende Zeiträume werden weniger übereinstimmend berichtet, weil sie vergessen werden.

Für die am weitesten zurückliegenden Jahre ist die Übereinstimmung in beiden Vergleichsrichtungen gering, was vor allem für die Arbeitslosigkeiten ohne Leistungsbezug (Inkonsistenzen vom Typ B) eher auf dem Alter beim Erleben als auf der Länge der seither verstrichenen Zeit beruht. Hier besteht auch während folgenschwerer und langer Phasen unfreiwilliger Erwerbslosigkeit oft noch keine Bezugsberechtigung für Leistungen der Bundesanstalt für Arbeit. Deshalb können auch lange und folgenschwere Phasen der Erwerbslosigkeit, die als Arbeitslosigkeitsphasen repräsentiert sind, ohne Leistungsbezug sein. Umgekehrt sind Leistungsbezugszeiträume in den frühen Jahren der Erwerbskarriere oft kurze, folgenarme und erwartete Übergangsphasen zwischen Ausbildung und Beruf oder in der Orientierungsphase und werden deshalb eher unterschlagen.

Die Veränderung der Übereinstimmung über den Zeitraum hinweg lässt sich also teilweise als eine Kombination aus Effekten der sozialen Erwünschtheit, des Vergessens und dem Alter beim Erleben deuten. Allerdings lässt sie sich nicht in allen Teilen erklären. Besonders die kurzzeitige deutliche Verbesserung der Übereinstimmung in beide Vergleichsrichtungen etwa 12 bis 13 Jahre vor der Befragung ist nicht einleuchtend aus den beiden unterschiedlichen Datenkonstruktionsprozessen erklärlich. Für die Inkonsistenzen vom Typ B kann außerdem die schlechte Übereinstimmung für das Jahr direkt vor der Befragung und den anschließenden u-förmigen Verlauf nicht mit sozialer Erwünschtheit erklärt werden; hier wäre eher im Gegenteil zu erwarten gewesen, dass in geringerem Umfang Zeiträume ohne Leistungsbezug als Arbeitslosigkeit berichtet werden. Hier wären genauere

Kenntnisse darüber erforderlich, wie Arbeitslosigkeit in der retrospektiven Lebensgeschichte eingeschätzt und bewertet wird.

Die Befunde betreffen die Arbeitslosigkeiten im Zeitraum Ende der 1970er bis Ende der 1990er Jahre für westdeutsche Männer und Frauen der Geburtsjahrgänge 1964 und 1971, die bei der Befragung ihre Sozialversicherungsnummer für die Datenverknüpfung zur Verfügung gestellt hatten. Die Generalisierung auf andere Gruppen und andere Zeiträume muss mit der gebotenen Vorsicht getätigt werden. Wie in Kapitel 5 dargestellt, haben schon Ostdeutsche aufgrund ihrer unterschiedlichen Sozialisation und Arbeitslosigkeitserfahrung eine andere Wahrnehmung von Arbeitslosigkeit als Westdeutsche, und Holst (2000) zufolge haben Ostdeutsche auch eine noch höhere Neigung, sich bei bestehender Bezugsberechtigung auch wirklich arbeitslos zu melden. Dies würde schon etwa bei der Verknüpfung von Befragtenangaben aus Ostdeutschland mit der LED zu anderen Befunden führen. Leider besteht keine Möglichkeit, den Registerdatenvergleich auf die LV-Ost auszudehnen, da diese nicht nach ihrer Sozialversicherungsnummer gefragt worden waren.

Auch dass die Befragten sich im jungen Erwachsenenalter befanden und nur auf die ersten 10 bis 15 Jahre ihres Erwerbslebens zurückblickten, schränkt die Generalisierungsmöglichkeiten ein. Wie sich Arbeitslosigkeit aus noch längerer Rückschau darstellt und zu Zeiträumen des Leistungsbezugs verhält, ist nur schwer abzusehen. Eine Befragung älterer Personen oder idealerweise eine Rekontaktierung und Neubefragung derselben Personen in 15 bis 20 Jahren wäre hier äußerst aufschlussreich und wird auch von den Entwicklungen des Arbeitsmarkts und des Beschäftigungssystems in diesem Zeitraum abhängen. Gerade im Hinblick auf das „Reminiszenzphänomen“ (siehe Abschnitt 2.6.2), das eine besondere Rolle für Ereignisse aus dem jungen Erwachsenenalter für die Konstitution des Selbst und der Lebensgeschichte nahe legt, wäre es interessant zu beobachten, wie sich die Erinnerung an diese Phase später im Lebenslauf verändert und wie dies die Übereinstimmung mit den Leistungsbezugszeiträumen beeinflusst. Zuletzt ist anzunehmen, dass diejenigen, die zum Zeitpunkt der Befragung ihre Sozialversicherungsnummer zur Verfügung stellen wollen und können, eine Auswahl von Personen darstellen, die ihre Erwerbstätigkeit ernst nehmen und sie in den Mittelpunkt stellen.

#### *Registerdatenvergleich als externe Validierung?*

Die Abweichungen treten unter der Bedingung auf, dass keinerlei Definition dessen gegeben wird, was unter Arbeitslosigkeit zu verstehen ist und somit die Befragten anhand des von ihnen selbst interpretierten Abruf-Hinweisreizes ans Rekonstruieren gehen. Würden die Befragten aufgefordert, Zeiträume zu erinnern, in denen sie Leistungen der Bundesanstalt für Arbeit bezogen haben, würde die Übereinstimmung unter Umständen höher ausfallen. Andererseits handelt es sich dabei sicherlich nicht um einen Begriff, mit dem Lebensabschnitte oder andere Einheiten des autobiografischen Gedächtnisses gebildet oder organisiert werden, sodass die Rekonstruktion diese Zeiträume als Bestandteil der Erwerbsgeschichte erheblich komplizierter würde. Darüber hinaus ist zu fragen, ob Arbeitslosigkeit im Lebenslauf so sinnvoll und angemessen erfasst werden könnten.

Der Vergleich hat gezeigt, dass Abweichungen zwischen den Datensätzen nicht pauschal als Fehler *beim Abbilden von Arbeitslosigkeit in den Befragungsdaten* gewertet werden können. Zwar kann bei den meisten Leistungsbezugszeiträumen davon ausgegangen werden,

dass die Hauptaktivität der Person im entsprechenden Zeitraum auch Arbeitslosigkeit ist und als solche in der Befragung berichtet werden sollte; dies ergibt sich aus den Voraussetzungen für Leistungsbezug und aus den Anforderungen an den Leistungsbezieher (aktives Bemühen um Arbeit, Stellensuche usw.). Andererseits gibt es zahlreiche gute Gründe dafür, Zeiträume unabhängig vom Leistungsbezug dann als Arbeitslosigkeit zu erfassen, wenn die Hauptaktivitäten der Befragten in einem bestimmten Zeitraum darin bestanden, die Rückkehr auf den Arbeitsmarkt anzustreben. Tatsächlich beklagen Arbeitsmarktforscher, dass über den Antragstellerstatus selektiv bestimmte Altersgruppen und Frauen nicht erfasst werden, die gemäß der ILO-Definition oder der Selbsteinschätzung zu den Arbeitslosen zu rechnen wären (Gallie & Vogler, 1994; McGinnity, 2001). Hier werden Aspekte von Arbeitslosigkeit durch die Befragung sogar angemessener abgebildet als durch die Leistungsbezugszeiträume, die solche Arbeitslosigkeiten nicht erfassen können. Insofern sind Inkonsistenzen vom Typ A eher als Fehler der Befragungsdaten zu werten. Allerdings gibt es auch hier Fälle, in denen die LV-West vorsieht, Zeiträume ungeachtet ihres Leistungsbezugsstatus einem anderen Zustand aus dem Zustandsraum zuzuordnen (Hillmert, Künster, Spengemann, & Mayer, 2004).

Oft lässt sich nur in einer Einzelfallprüfung differenziert entscheiden, welche Daten die angemessenere Abbildung darstellen. Als Beispiel möchte ich die Einzelfälle 5\_1, 5\_2 und 5\_3 erörtern, die alle einen Leistungsbezugszeitraum in eine andere Aktivität umdefiniert haben. Bei Person 5\_1 wäre es angemessener, die Aushilfskellnerei ins Nebenerwerbsmodul aufzunehmen und hier als Hauptaktivität Arbeitslosigkeit einzutragen. Bei Person 5\_2 ist die Zuordnung schwierig; am besten würden hier Arbeitslosigkeit und Ausbildung parallel bestehen bleiben, sodass weder eine Phase der Arbeitslosigkeit noch eine Bildungsperiode verloren geht. Bei Person 5\_3 hingegen scheint es angemessen, den Zeitraum als Ausbildung darzustellen – die Tatsache, dass es sich um eine arbeitsamtfinanzierte Ausbildung handelte, ist eine wichtige Detailinformation, würde aber nicht die Abbildung als Arbeitslosigkeit rechtfertigen. Gerade arbeitsamtfinanzierte Ausbildungen können mehrere Jahre dauern und deshalb besonders gravierende Inkonsistenzen zwischen LED und LV-West ergeben. Diese Abweichungen stellen keine Fehler in den Befragungsdaten dar, sondern sind den Nichtvergleichbarkeiten der „Verwaltungstheorie“ der Bundesanstalt für Arbeit und dem Konzept der LV-West von Arbeitslosigkeit geschuldet.

Diese Überlegungen machen deutlich, dass die Frage der Validität von Arbeitslosigkeitsdaten auf zwei Ebenen gestellt und beantwortet werden muss:

- (1) Werden durch eine Datenkonstruktionsprozedur alle Arbeitslosigkeiten gemäß der verwendeten Definition auch tatsächlich erfasst?
- (2) Ist der verwendete Indikator ein valider bzw. angemessener Zugang zu allen relevanten Facetten und Aspekten des sozialwissenschaftlichen Tatbestands „Arbeitslosigkeit“?

Wenn die Datenkonstruktionsprozesse und die zu erhebenden Konzepte in wesentlichen Punkten voneinander abweichen, ist es sinnvoll, den Vergleich nicht einseitig als Validierung des einen Datensatzes am anderen zu führen. Stattdessen kann die Abbildung ein und desselben Konzepts in den beiden Datensätzen unter den Gesichtspunkten einer *differenziellen* oder *komplementären Validität im Hinblick auf den zu erfassenden Tatbestand* verglichen werden. Dem liegt ein Konzept von Validität als Angemessenheit zu Grunde (Rohwer & Pötter, 2002), das gerade bei Konzepten wie Arbeitslosigkeit Sinn macht, die die Interessen der Wissenschaftler und Konventionen des Wissenschafts- bzw. Verwaltungssystems widerspiegeln. In

jedem Fall ist es unverzichtbar, die Abweichungen unter genauer Kenntnis und Berücksichtigung beider Datenkonstruktionsprozesse sowie der konzeptuellen und theoretischen Hintergründe *beider* Datensätze zu ermitteln und zu interpretieren<sup>11</sup>.

---

<sup>11</sup> Nicht hoch genug einzuschätzen ist dabei der Wert von Personen, die einen oder beide Datensätze intim kennen, um deren „Wissen in Hinterköpfen“ (Bick & Müller, 1980) anzuzapfen – sowohl was den offiziellen Datenkonstruktionsprozess als auch die pragmatischen Realitäten des Datenkonstruktionsprozesses betrifft (beispielsweise welche Fragen problematisch waren, wo Angaben unvollständig sind). Beispielsweise konnte der Sachverständige für die LV, Ralf Künstler, mitteilen, dass das im entsprechenden Datenfile abgelegte Interviewdatum nicht immer dem Datum der Erhebung der Erwerbsgeschichte entsprach. Wenn Befragungen aus Zeitgründen in zwei oder drei Etappen durchgeführt werden mussten, war die Erwerbsgeschichte jeweils im ersten Block, und die Datumsangaben der ersten Befragungsblöcke waren in einem anderen Datensatz abgelegt. Ohne dieses „Insiderwissen“ hätte ich für mindestens 20 Personen ein falsches Interviewdatum bei der Berechnung der retrospektiven Distanz zu Grunde gelegt.



## Kapitel 7

### Diskussion und Ausblick: Zusammenfassung und Implikationen

Ich möchte nun zusammenfassen, was die drei empirischen Analysen über die Auswirkungen von autobiografischen Erinnerungsprozessen auf die Daten- und Datenfehlergenerierung in retrospektiven standardisierten Lebensverlaufsbefragungen ergeben haben. Danach möchte ich die Implikationen diskutieren, die sich daraus für die Qualitätsabschätzung und -steigerung von retrospektiven Befragungsdaten ableiten lassen. Die Diskussion wird sich zum einen mit der Vermeidung und Korrektur von Gedächtnisfehlern in Erhebung und Edition befassen, zum anderen mit der Art, wie Datenfehler und besonders Erinnerungsfehler in Befragungen beschrieben, interpretiert und dokumentiert werden können.

#### **7.1 Autobiografisches Erinnern als Fehlereinfluss in standardisierten retrospektiven Befragungen**

##### *Rekonstruierte Lebensläufe: Veränderungen in der Retrospektive*

In dieser Arbeit habe ich untersucht, in welcher Weise die Neigung der autobiografischen Informationsverarbeitung, durch selektives Vergessen und Umdeuten kontinuierlichere und konventionellere Lebensverläufe zu rekonstruieren, zu Fehlern bei der retrospektiven Erhebung sozialwissenschaftlicher Längsschnittdaten in standardisierten Befragungen führt. Dazu habe ich drei empirische Studien mit den retrospektiven lebensgeschichtlichen Ereignisdaten aus der *Deutschen Lebensverlaufsstudie* (LV) durchgeführt. Zum einen habe ich die Reliabilität der Angaben aus einer der ostdeutschen Teilstudien zu Kindern und zur Erwerbsbiografie geprüft. Zum anderen habe ich die Angaben über Arbeitslosigkeitsphasen aus einer westdeutschen Teilstudie der LV mit den Daten der Leistungsempfängerdatei der Bundesanstalt für Arbeit (BA) verglichen und damit deren Validität abgeschätzt.

In allen drei Studien führen Prozesse des autobiografischen Erinnerns zu einer spezifischen Gedächtnisfehlerkomponente in den Daten, die die Reliabilität und die Validität in erheblichem Maße mit beeinflusst. Ereignisse, die nicht der aktuellen und normativen Entwicklungslogik des Selbst entsprechen, werden eher vergessen, und an deren Stelle wird der „Standard“ erinnert, den diese normativen und individuellen Erwartungen setzen. Als Resultat fallen Angaben aus späteren Befragungen einfacher und konventioneller aus als die aus früheren Befragungen, und Befragungsdaten fallen einfacher und konventioneller aus als Registerdaten.

Für die Angaben zu den Kindern führen Gedächtnisprozesse vor allem dazu, dass nicht-lebliche und bereits gestorbene Kinder aus der längeren retrospektiven Distanz wesentlich seltener erwähnt werden. Außerdem werden die Auszüge der Kinder aus dem elterlichen Haushalt häufig inkonsistent datiert und dabei in die Nähe des schematischen „Erwartungswerts“ um 21 Jahre gerückt.

Für die Rekonstruktion der Nachwende-Erwerbssequenzen als Abfolge von Erwerbstätigkeiten, Arbeitslosigkeiten und Nichterwerbstätigkeiten führen Gedächtnisprozesse dazu, dass aus längerer retrospektiver Distanz wenig bedeutsame und folgenschwere Veränderungen, die nicht als Übergang in einen neuen Lebensabschnitt repräsentiert sind, häufig eingebnet werden. Die Erinnerung an Arbeitslosigkeit und Nichterwerbstätigkeit ist ebenfalls instabil. Episoden werden sowohl unterschlagen als auch hinzugefügt. Die dadurch entstehenden zeitlichen Lücken werden meistens durch Strecken der Standardaktivität gefüllt; dies ist in der Regel Erwerbstätigkeit.

Im Vergleich mit der Leistungsempfängerdatei führen Gedächtnisprozesse vor allem dazu, dass die Befragten kurze, in „Serie“ auftretende und untypisch lange Leistungsbezugszeiträume, während derer andere Aktivitäten im Vordergrund standen, nicht angeben. An deren Stelle treten meist Erwerbstätigkeiten, teilweise wurden die Episoden auch in Nichterwerbstätigkeiten umgedeutet. Außerdem werden zahlreiche und auch ausgedehnte Zeiträume als Arbeitslosigkeit angegeben, in denen keinerlei Leistungen bezogen wurden, besonders aus der Zeit zwischen Schule, Ausbildung, Wehrdienst und den ersten Erwerbstätigkeiten. Da die Definition von Arbeitslosigkeit in den Befragungs- und Registerdaten nicht genau übereinstimmt, kann man hier nicht uneingeschränkt von Gedächtnisfehlern sprechen. Der Vergleich muss vielmehr in einem größeren Rahmen der Erfassung von Arbeitslosigkeit gesehen werden, in dem man die Abbildung ein und desselben Konzepts in den beiden Datensätzen unter den Gesichtspunkten einer *differenziellen* oder *komplementären Validität im Hinblick auf den zu erfassenden Tatbestand* vergleicht.

Da Lebensverläufe in Einklang mit der individuellen Gedächtnisorganisation rekonstruiert werden, treten gedächtnisverursachte Abweichungen stets in alle Richtungen auf, das heißt, dass durchaus auch Lebensläufe retrospektiv wechselhafter und komplexer rekonstruiert werden, indem Übergänge oder Episoden nachträglich eingefügt wurden. Die Gesamt-tendenz zum Vereinfachen und Konventionalisieren entsteht, weil diese Fälle seltener auftreten.

Die Neigung zur Konventionalisierung der rekonstruierten Sequenz muss vor dem Hintergrund der jeweiligen Erinnerungskontexte gesehen werden, die die normativen Erwartungen über Lebensverläufe mitbestimmen und den Standard setzen, an den die Angaben angenähert werden. So lässt sich die Zunahme von Arbeitslosigkeitsepisoden aus der längeren retrospektiven Distanz, die beim Vergleich der Wiederholungsbefragung in Ostdeutschland auftrat, als eine Folge einer veränderten Normalbiografie deuten; ich werde weiter unten noch näher darauf eingehen.

#### *Ursachen, Folgen und Edition*

Der Gedächtnisfehler fällt für Männer und Frauen und für unterschiedliche Altersgruppen verschieden aus, wobei aber Inhalte und Fehleraspekte differenziert betrachtet werden müssen. Männer machen in fast jeder Hinsicht fehlerhaftere Angaben zu ihren Kindern und deren Auszügen aus dem Elternhaus; in der Erwerbsgeschichte finden sich hingegen keine konsistenten Unterschiede über alle Inhalte, Informationsdimensionen und Fehleraspekte hinweg. Dies lässt sich als Ergebnis der besonderen Situation der ostdeutschen Befragten deuten, da Frauen und Männer aufgrund ihrer beruflichen Sozialisation in der DDR ähnliche Erwartungen und Vorstellungen über ihre Erwerbstätigkeit haben. Hier sind weiterführende Analysen

und spezifischere Hypothesen nötig, um die Komplexitäten und Bedingungen der Unterschiede besser zu verstehen.

Ähnlich stellt sich das Bild für die befragten Kohorten dar. Sowohl bei den Angaben zu Kindern als auch bei den Angaben zu Erwerbsverläufen machen teilweise die Angehörigen der älteren Kohorten mehr oder größere Fehler, teilweise die der jüngeren oder der mittleren Kohorten. Auch dieser Umstand muss vor dem Hintergrund betrachtet werden, dass die Rekonstruktion von Erwerbsverläufen im Transformationszeitraum für alle Altersgruppen die Rekonstruktion einer Zeit voller radikaler Veränderungen und Unsicherheiten im Erwerbsleben darstellt. Auch hier wären weiterführende Analysen mit spezifischeren Unterschiedshypothesen aufschlussreich.

Da in den Analysen die Abweichungen und Fehler im Vordergrund standen, ist es wichtig herauszustellen, dass sehr viele der Befragten konsistente Angaben machen. Gerade die zentralen lebensgeschichtlichen Informationen werden sehr stabil rekonstruiert, und wenn ein bestimmtes Ereignis nie aufgetreten ist – zum Beispiel wenn eine Person nie arbeitslos war oder keine Kinder hat –, sind die Angaben ausgesprochen zuverlässig. Vor allem wechselhafte Verläufe, die nicht den individuellen und allgemeinen Normen über Lebensverläufe entsprechen, werden zeitlich instabil und inkorrekt rekonstruiert. Dies stellt natürlich eine Herausforderung dar, wenn Lebensläufe insgesamt weniger standardisiert verlaufen. Zum Beispiel belegt Hullen (1998), dass Auszüge aus dem Elternhaus immer seltener mit Heirat und Familiengründung zusammenfallen, wodurch sie zunehmend verschwommen und uneindeutiger in der Zeit lokalisierbar werden. Für das Erwerbsleben beschreiben Schlagworte wie „Destandardisierung“ und „Pluralisierung“, dass institutionell nur schwach geregelte Verläufe mit Teilzeitstellen, Familienarbeitsphasen, Zeitarbeit, befristeten Verträgen und Weiterbildungen usw. zunehmend häufiger auftreten und verstärkt auch Männer und spätere Lebensabschnitte betreffen (Berger & Sopp, 1992; Kohli, 1985; Mayer, 1998). Solche Biografien stellen generell eine Herausforderung für das Sequenzkonzept des Lebenslaufs dar. In der LV wird deshalb seit der ersten Erhebung zusätzlich zur Hauptsequenz ein weniger rigider Nebenpfad von Nebentätigkeiten und Weiterbildungen usw. erhoben, und die Teilstudie *Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland (LV-West)* lässt bereits in der Hauptsequenz die Gleichzeitigkeit von Erwerbstätigkeiten und (Vollzeit-)Ausbildungen zu, um betriebliche und berufs begleitende Ausbildungen angemessener abbilden zu können.

Durch den Gedächtnisfehler werden auch die Ergebnisse deskriptiver und explanatorischer Analysen beeinflusst. Abweichungen sind besonders zu erwarten, wenn in den abhängigen oder unabhängigen Variablen zeitliche oder Verlaufsaspekte der Lebensläufe wie Inzidenz, Dauer, Timing oder Reihenfolge von Ereignissen kodiert sind und beispielsweise die Inzidenz und Dauer früherer Zustände zur Erklärung der Inzidenz und Dauer späterer Zustände herangezogen werden. Dies steht auch in Übereinstimmung mit den Analysen von Paull (2002) und Pierret (2001). Welchen Effekt Gedächtnisfehler im konkreten Fall genau auf die Stichprobenkennwerte oder Modellparameter haben wird, ist schwer vorherzusagen. Auf den Stellenwert solcher Vergleiche werde ich deshalb weiter unten noch gesondert eingehen.

Die deutlich höhere Übereinstimmung der edierten Daten belegt, dass durch das arbeits- und kostenintensive Vorgehen der Einzelfalldition viele der Gedächtnisfehler korrigiert werden konnten. Nahezu alle der nicht angegebenen leiblichen und ein großer Teil der nichtleiblichen Kinder wurden in der Edition nachgetragen, und für die Erwerbsverläufe haben sich nahezu alle Abweichungen in jeder Hinsicht deutlich reduziert.



Leider profitieren die Informationen, die auch schon bei der Erhebung inkonsistenter waren, auch von der Edition in geringerem Maß. Dies betrifft besonders die innerbetrieblichen Wechsel und die Verwechslung von Arbeitslosigkeiten und Nichterwerbstätigkeiten. Da innerbetriebliche Wechsel seltener einen neuen Lebensabschnitt markieren, werden sie erstens weniger stabil rekonstruiert und zweitens ruft es selten Widersprüche innerhalb der Angaben hervor, wenn sie eingegeben werden. Ebenso können sich Arbeitslosigkeiten und Nichterwerbstätigkeiten in ihren Aktivitäten und Rollen ähneln, sodass das eine in das andere umgedeutet werden kann, ohne dass dies zu Widersprüchen in den Angaben führt. Gerade diese Information sollte also schon in der Erhebung möglichst korrekt erfasst werden.

#### *Arbeitslosigkeit: Repräsentation, Rekonstruktion und Leistungsbezug*

Identität, soziale Rollen und Status, Teilhabe an gesellschaftlichen Angeboten, Ziele, Motive und Lebenszufriedenheit hängen zumindest in industrialisierten Gesellschaften in starkem Maß mit dem Erwerbsbereich zusammen, wie vor allem Studien über die psychosozialen Folgen des Verlustes von Erwerbsarbeit belegen (Gallie & Russell, 1998; Jahoda, 1982; Wacker, 1976). Die Erwerbsgeschichte eines Menschen ist deshalb als zentraler Bereich des Selbst in hohem Maße organisiert und als ein den subjektiven und normativen Entwicklungslogiken angepasstes Narrativ repräsentiert, das die Gegenwart als Folge der Vergangenheit etabliert.

Die beiden Vergleiche zur Erwerbsgeschichte verraten uns Näheres über die Voraussetzungen dafür, dass ein Zeitraum innerhalb dieses Sinngefüges als Arbeitslosigkeitsepisode oder Lebensabschnitt in Arbeitslosigkeit repräsentiert ist. Sich arbeitslos gemeldet zu haben, Leistungen zu beziehen und die vom Arbeitsamt geforderten Suchaktivitäten an den Tag zu legen und nachzuweisen, ist offenbar weder eine notwendige noch eine hinreichende Voraussetzung dafür. Arbeitslosigkeit im Lebenslauf definiert sich negativ über die unerwünschte Abwesenheit von Erwerbsarbeit (Elias, 1991; Mutz, 1996). Trotzdem können Arbeitslosigkeitsphasen zu stabil repräsentierten Lebensabschnitten werden – etwa als Brüche in der Erwerbsbiografie, als Wendepunkte für eine Neuorientierung des Lebens, als überwundene Stolpersteine bei der (erfolgreichen) Rückkehr in die Erwerbsarbeit oder als Anfangs- und Endpunkte eines sozialen Abstiegs. Diese Einschätzung kann sich mit der Zeit verändern, wenn sich die Interpretations- und Bewertungsgrundlagen für Arbeitslosigkeitsabschnitte im Lebenslauf ändern. Dies ist offenbar in Ostdeutschland zwischen Anfang und Ende der 1990er Jahre geschehen. Zahlreiche Autoren belegen, dass das individuelle, autobiografische Gedächtnis und das historische Gedächtnis in enger Beziehung stehen (Brown, 1990; Brown, Shevell, & Rips, 1986; Lieury, Richer, & Weeger, 1978) und ordnen die individuelle biografische Konstruktion in kollektive Gedächtnisprozesse ein (Halbwachs, 1985; Middleton & Edwards, 1990). Durch die langfristige Beschäftigungs- und Strukturkrise in Ostdeutschland wird Arbeitslosigkeit möglicherweise zu einer kollektiven historischen Erfahrung und ist als solche in geringerem Maße aus der individuellen Ziel- und Handlungslogik der Biografie heraus begründungsbedürftig.

#### *Die Rolle der verstrichenen Zeit: Je länger die Rückschau, desto fehlerhafter die Rekonstruktion?*

In allen drei Analysen zeigt sich, dass die retrospektiven Rekonstruktionen immer vereinfachte und konventionellere Versionen der Realität darstellen, ganz gleich wie lange der Zeitraum zurückliegt, über den berichtet wird. Darüber hinaus gab es Anhaltspunkte, dass diese Tendenz für länger zurückliegende Zeiträume gravierender ausfällt, sodass ein Ereignis umso

wahrscheinlicher vergessen und umso abweichender rekonstruiert wird, je länger es zurückliegt. Dies muss allerdings differenziert für verschiedene Angaben und verschiedene Aspekte des Fehlers betrachtet werden. Beispielsweise nimmt, wie in Abschnitt 4.3.3 berichtet, das *Irrtumsrisiko* bei der Datierung von Auszügen über die Zeit nicht zu, die *Fehlergröße* aber nimmt linear zu. Da es sich um den Vergleich von Wiederholungsbefragungen handelt, steht dies für eine lineare Zunahme des Irrtumsrisikos und eine exponentielle Zunahme der Fehlergröße.

In einer retrospektiven Befragung ist die Zeitspanne zwischen Erleben und Rekonstruktion stets mit einer Vielzahl von Einflussfaktoren assoziiert, die die autobiografische Rekonstruktion beeinflussen. In den hier durchgeführten Analysen fanden sich Effekte des Alters beim Erleben eines Ereignisses, Auswirkungen historischer Veränderungen in den tatsächlichen Tatbeständen und der damit verbundenen Veränderungen der Repräsentationen und Kategorieninterpretationen sowie zeitabhängig unterschiedlich ausfallende Effekte der sozialen Erwünschtheit<sup>1</sup>.

Manche dieser Einflussfaktoren wirkten tatsächlich dahingehend, dass die länger zurückliegenden Ereignisse schlechter erinnert wurden, so etwa die Tatsache, dass Arbeitslosigkeiten in der Phase des Berufseintritts deutlich schlechter mit Leistungsbezugszeiträumen übereinstimmten. Beim Vergleich der ostdeutschen Angaben zu den Erwerbsverläufen besteht aber durchaus die Möglichkeit, dass sozialwissenschaftliche Tatbestände wie „Arbeitslosigkeit“ oder „Erziehungsurlaub“ durch die längere Vertrautheit mit dem System, aus dem diese Konzepte stammen, in der zweiten Befragung eher im Sinne des Wissenschaftlers angewendet werden. Auch war Arbeitslosigkeit zum Zeitpunkt der zweiten Befragung eine kollektive „Standarderfahrung“ geworden, sodass Selbstdarstellungsmotive im Sinne von sozialer Erwünschtheit vermutlich schwächer ausgeprägt sind. Leider fehlt in dieser Reliabilitätsstudie der Bezug zu einem externen Standard, der ermöglichen würde zu entscheiden, welche Angabe jeweils die angemessenere darstellt. Im Registerdatenvergleich zumindest war die Übereinstimmung für die Zeit unmittelbar vor der Befragung vergleichsweise schlecht, weil die Befragten teilweise nicht zugeben wollten, arbeitslos zu sein<sup>2</sup>.

## 7.2 Implikationen für die Erhebung und Edition retrospektiver Lebensverlaufsdaten

Da retrospektive Erhebungen eine Reihe von Vorteilen gegenüber Panelbefragungen haben und da darüber hinaus retrospektive Elemente auch in Panelbefragungen verbreitet sind, ist es wichtig, bei der Erhebung und der Edition gezielt die Vermeidung und Korrektur ihrer spezifischen Schwächen zu betreiben.

<sup>1</sup> Effekte des Alters in dem Sinne, dass Ältere sich aufgrund nachlassender kognitiver Effizienz schlechter erinnern, traten in dieser Arbeit keine auf. Allerdings waren auch die ältesten Befragten zum Befragungszeitpunkt höchstens 59 Jahre alt und damit in einem Alter, in dem merkliche kognitive Beeinträchtigungen noch nicht zu erwarten sind.

<sup>2</sup> Abweichungen aufgrund von sozialer Erwünschtheit können nicht als ein Fehler bei der *Erinnerung* behandelt werden, da es sich um absichtsvolles Verschweigen handelt. Bei der Einschätzung spezifischer Stärken und Schwächen von retrospektiven und auf die Gegenwart bezogenen Daten muss man allerdings auch die zeitabhängig unterschiedlichen Fehlerkomponenten aus anderen Bereichen des Gesamtbefragungsfehlers berücksichtigen, um Vergessenskurven zu begründen und die Güte von Angaben für frühere und spätere Zeiträume einzuschätzen.

Viele Informationsdimensionen der Ereignisdaten bleiben fehlerfrei, wenn die Episodenstruktur der Lebensverläufe korrekt rekonstruiert wird. Dies gilt besonders für Informationen über Inzidenz, Frequenz und Dauer von Arbeitslosigkeit und Nichterwerbstätigkeit. Auch die Information über das Timing von Veränderungen und Übergängen ist konsistenter, wenn Episoden und Übergänge äquivalent rekonstruiert werden; vor allem treten Abweichungen dann weitestgehend symmetrisch auf. Deutliche und umfassende Steigerungen der Datengüte können also erzielt werden, wenn die korrekte Rekonstruktion der Episodenstruktur eines Lebens in der Befragung gelingt oder nachträglich in der Edition hergestellt wird. Ersteres ist dabei zu bevorzugen, denn wenn nachträglich Episoden eingefügt, gestrichen, geteilt oder fusioniert werden, ist die Detailinformation zu den ursprünglichen Episoden und Übergängen nicht ohne weiteres auf die veränderte Struktur übertragbar.

#### *Verbesserung der Rekonstruktionen bei der Erhebung*

Die zeitstabile (reliable) und den Zielen der LV angemessene (valide) Rekonstruktion von Episoden und Übergängen kann optimiert werden, indem dem Befragten ein genaueres Verständnis dafür vermittelt wird, was von ihm verlangt wird. Spezifischere Abruf-Hinweise, die die Schwierigkeiten bei der Rekonstruktion bestimmter Angaben berücksichtigen, könnten helfen, uneindeutige Episoden angemessener zuzuordnen und Übergänge und Episoden, die nicht Bestandteil des Lebensnarrativs sind, doch zugänglich zu machen.

Ich möchte dies mit einigen Beispielen aus den drei empirischen Analysen verdeutlichen. Die Ergebnisse aus Kapitel 4 belegen, dass das scheinbar eindeutige Ereignis „Auszug eines Kindes“ reichlich Interpretationsspielraum bietet. Dieser wird durch die Konventionalisierungseigung des autobiografischen Gedächtnisses zu späteren Zeitpunkten tendenziell anders genutzt. Auf einem Tonbandmitschnitt der ersten Befragung besprachen sich der Befragte und seine Frau etwas ratlos darüber, wann ihr Sohn denn ausgezogen sei – als er zum Militär ging oder als er über ein Jahr danach auch seine Möbel aus dem elterlichen Haushalt in eine eigene Wohnung überführte. Den Befragten war also durchaus klar, dass hier eine Entscheidung für eines von mehreren Ereignissen getroffen werden musste, die Frage selbst bot ihnen allerdings keinen Anhaltspunkt, der ihnen ermöglicht hätte, im Sinne der Wissenschaftler zu entscheiden.

Als zweites Beispiel möchte ich die *Erhebung der innerbetrieblichen Wechsel* in der LV-Ost anführen, die in der zweiten Befragung sehr inkonsistent erinnert worden waren. Ein Befragter, der eine Erwerbstätigkeit angegeben hat, wird durch die Frage „Bis wann waren Sie in dieser Stellung tätig?“ veranlasst, die nächste Veränderung zu rekonstruieren<sup>3</sup>. Was sich genau bei einem Wechsel verändert hat, wird als Detailinformation zum Übergang erst erfragt, *nachdem* dieser rekonstruiert wurde. Der Befragte weiß also zunächst gar nicht, welche Veränderungen er als Wechsel berichten *soll*, und gibt deshalb die erste in der retrospektiven Wahrnehmung seiner Lebensgeschichte relevante Veränderung in den erwerbsbezogenen sozialen Rollen und Aktivitäten an, also den Übergang in den nächsten Lebensabschnitt. Dies ist nicht notwendigerweise die erste Veränderung im Sinne der LV, und es ist nicht notwendigerweise bei einer

<sup>3</sup> Die LV sieht vor, dass innerhalb Phasen kontinuierlicher Erwerbstätigkeit Veränderungen des Einkommens, der Arbeitszeit, der beruflichen Stellung oder des Betriebs erfasst werden sollen sowie Übergänge in Arbeitslosigkeit oder Nichterwerbstätigkeit.

anderen Rekonstruktion derselbe Übergang. Ein spezifischerer Abruf-Hinweisreiz könnte auch die schwach repräsentierten Übergänge aktivieren – etwa indem dem Befragten erläutert wird, was alles einen Übergang im Sinne der LV darstellt.

Als drittes Beispiel möchte ich die *inkonsistente Einordnung uneindeutiger Zustände* anführen, wenn etwa Zeiträume mit Leistungsbezug als Erwerbstätigkeiten eingestuft werden, weil der Befragte stundenweise gekellnert hatte, oder als Zeit in Mutterschutz eingeordnet werden, weil die Befragte in dieser Zeit einen Säugling zu versorgen hatte. Eine Definition dessen, was unter Arbeitslosigkeit verstanden werden soll, würde einen spezifischeren Abruf-Hinweisreiz darstellen, anhand dessen die Rekonstruktion jedes Mal im Sinne des Wissenschaftlers vorgenommen werden kann<sup>4</sup>. Beispielsweise berichten Norwood und Tanur (1994), dass Arbeitslosigkeit gemäß der Definition der *International Labor Organisation* (ILO) besser erfasst wurde, wenn jeder der Aspekte (u.a. aktive Suche nach Arbeit, Arbeitsfähigkeit und Verfügbarkeit) in einer eigenen Frage „abgeklopft“ wurde. Dies ist in retrospektiven Befragungen nicht durchführbar, belegt aber die Effizienz, mit der spezifische Abruf-Hinweisreize den Befragten dazu bringen können, sein „privates“ Erwerbsnarrativ zu verlassen und Arbeitslosigkeit im Sinne des Wissenschaftlers zu betrachten.

#### *Über den individuellen Erinnerungsfehler hinaus: Kooperative Befragungen*

Ein partielles, aufs autobiografische Erinnern als individuellen Prozess beschränktes Datengenerierungsmodell hat seine Grenzen und muss für eine weitergehende Aufklärung der Datenfehlergenerierung in umfassendere Modelle der Befragungssituation als soziale Interaktionssituation eingebettet werden (Sudman & Bradburn, 1982; Sudman, Bradburn, & Schwarz, 1996). Zahlreiche Befunde belegen, dass sowohl der Interviewer als auch der Befragte eine Befragung tatsächlich als eine spezielle Form der Unterhaltung betrachten und sich entsprechend verhalten (Conrad & Schober, 2000; Houtkoop-Steenstra, 2000; Sudman & Bradburn, 1982; Sudman u.a., 1996). Hier ergeben sich Anknüpfungspunkte an die narrativen Aspekte autobiografischen Erinnerns, das auch als eine Form des sozialen Geschichtenerzählens begriffen wird (Gergen & Gergen, 1988; Robinson & Taylor, 1998), die Geschichten durch die Kommunikationssituation mitgestaltet werden.

Der Interviewer ist in dieser Situation der Stellvertreter des Wissenschaftlers und Vermittler seiner Interessen. Er benötigt deshalb ein gutes Verständnis der sozialwissenschaftlichen Konzepte und muss außerdem wissen, wie Information in retrospektiven Befragungen verarbeitet und wie in diesem Rahmen Fehler produziert werden und welches die relevanten Kontextfaktoren sind (z.B. ein instabiler Arbeitsmarkt). So kann er dem Befragten beim Erkennen von Uneindeutigkeiten und beim Auflösen derselben im Sinne des Wissenschaftlers behilflich sein.

Diese Kompetenzerweiterung bringt eine Lockerung der Standardisierung des Interviewerverhaltens mit sich. Allerdings ist bekannt, dass Interviewer ohnehin von dieser Standardisierung abweichen, um den Kommunikationsfluss aufrechtzuerhalten (Brückner, 1995; Houtkoop-Steenstra, 2000; Schaeffer & Maynard, 1996). Im oben genannten Beispiel, in

<sup>4</sup> Dies wäre in der LV nicht unbedingt „Zeiträume mit Leistungsbezug“. Wie ich in Kapitel 6 ausgeführt habe, gibt es zahlreiche gute Gründe, auch Zeiträume ohne Leistungsbezug als Arbeitslosigkeiten zu berücksichtigen und umgekehrt Leistungsbezugszeiträume nicht als Arbeitslosigkeiten zu betrachten.

dem sich das Ehepaar nicht schlüssig war, welches Ereignis den Auszug ihres Sohnes darstellte, fiel die Entscheidung mit „Unterstützung“ des Interviewers, der schnell vorankommen wollte und das erste, probeweise geäußerte Datum aufnotierte und die nächste Frage stellte. Insofern scheint es sinnvoll, dafür zu sorgen, dass die Interviewer eine angemessene Entscheidung im Sinne des Wissenschaftlers treffen können.

In einem solchen erweiterten Rahmen werden Interviewer und Befragter als Partner gesehen, die sich die konzeptuelle „Brille“ des Sozialwissenschaftlers für die Dauer der Befragung aufsetzen und in seinem Sinne Entscheidungen über Zuordnungen von Zeiträumen und Rekonstruktion von Übergängen treffen. Solche Konzepte der Befragungssituation werden als „kooperatives Interviewen“, „befragtenorientiertes Befragungsdesign (*recipient design*)“ oder „konversationales Interviewen“ bezeichnet und ihre Effektivität auch in experimentellen und Beobachtungsstudien untersucht (Conrad & Schober, 2000; Houtkoop-Steenstra, 2000; Schaeffer & Maynard, 1996; Suessbrick, Schober, & Conrad, 2000). Zwar betonen die Autoren, dass die Kostenseite (Interviewertraining, verlängerte Interviews, Risiko von suggestiven Beeinflussungen durch den Interviewer u.a.) noch nicht genug untersucht wurde. Insgesamt jedoch sind die Auswirkungen sowohl auf Datengüte als auch auf die Zufriedenheit von Befragten und Interviewer ermutigend.

#### *Erhebung und Edition – Verteilung der Aufgaben*

Eine Kosteneinschätzung methodischer Innovationen bei der Erhebung muss auch berücksichtigen, ob dadurch der Editionsbedarf reduziert werden kann. Hillmert (2002) berichtet, dass in der Erhebung *Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland (LV-West)* für die letzte und umfangreichste Editionsphase, die so genannte Endedition, etwa 5.000 Arbeitsstunden aufgewendet wurden. Für die Befragung der westdeutschen Geburtsjahrgänge 1929–31, 1939–41 und 1949–51 schätzen Mayer und Brückner (1989), dass die Edition noch einmal so viel Zeit und Ressourcen in Anspruch genommen hatte wie die eigentliche Erhebung selbst. Jede methodische Innovation, die die Datenqualität bereits bei der Erhebung steigert und damit den Editionsbedarf senkt, würde eher zu einer Umverteilung als zu einer Neuentstehung von Kosten führen.

In diesem Rahmen stellt sich die Frage, ob ein Datenaufzeichnungsinstrument wie die in Abschnitt 3.4.4 beschriebenen biografischen Kalender den Erhebungsprozess sinnvoll unterstützen kann. Solche Instrumente fördern die Güte retrospektiver Daten unter anderem dadurch, dass sie individuelle Kontexte zugänglich machen und dass die Kommunikation zwischen Interviewer und Befragtem unterstützt wird. Sie verlegen auch einen zentralen Aspekt der Editionsarbeit – die Herstellung eines konsistenten und plausiblen Gesamtlebenslaufs in mehreren aufeinander bezogenen Bereichen – in die Interviewsituation, da die grafische Aufbereitung der Angaben die Entdeckung und Bereinigung von Inkonsistenzen schon während der Datenerhebung ermöglicht und erfordert.

Eine derartige Verlagerung von Editionsarbeit hätte auch eine Verlagerung von Kompetenzen zur Folge. Die Editeure der LV waren in der Regel Studenten oder Absolventen der Soziologie oder Sozialwissenschaften, wurden intensiv geschult und standen in engem Kontakt mit den Sozialwissenschaftlern selbst. Ihnen standen viel Zeit und umfangreiche Unterlagen zur Verfügung und sie setzten sich in der Regel mit jedem Einzelfall als Gesamtlebensverlauf auseinander. Die Interviewer hingegen waren Honorarkräfte des Befragungsinstituts infas, die in

einer zweitägigen Schulung umfassend mit der Anwendung des CATI-Programms vertraut gemacht worden waren und kontinuierlich supervidiert und beraten wurden, aber sie hatten nie die Gesamtverläufe vor Augen und waren an die Struktur des CATI gebunden (infas, 1999).

Momentan ist noch nicht vollständig bekannt, wie biografische Kalender ihre Wirkung entfalten und welchen Rekonstruktionsherausforderungen sie besonders gut vorbeugen können. Als bei der Vorbereitung der ersten Erhebung der LV in einer Pilotstudie mit einer Lebensverlaufsmatrix experimentiert wurde, waren die Resultate gemischt (Papastefanou, 1980; Tölke, 1980); es schien, als ob das Instrument zwar den Schwächen des autobiografischen Erinnerns entgegenwirken konnte, dabei aber zugleich durch die Rigidität des Zeitrasters dessen Stärken an der Entfaltung hinderte. Zurzeit wird am Max-Planck-Institut für Bildungsforschung an der Entwicklung eines computerunterstützten Kalenderinstruments gearbeitet, das diese Schwierigkeiten überwinden und sowohl Erhebung als auch Edition unterstützen und kosteneffektiver gestalten soll.

### 7.3 Implikationen für die Fehlerforschung

In mehrfacher Hinsicht würden Fehlerforscher und Datenanwender davon profitieren, wenn Validierungsstudien häufiger durchgeführt sowie detaillierter und zugänglicher berichtet würden. Datennutzer mit spezifischen Fragestellungen könnten sich darüber informieren, wie sehr und in welcher Weise die sie interessierenden Variablen durch Erinnerungsfehler beeinflusst sind. Außerdem könnten Validierungsstudien direkter miteinander verglichen werden, um kumulativ das Wissen über die Phänomene und Prozesse autobiografischen Erinnerns bei der retrospektiven Verlaufsdatenerhebung zu einem differenzierten und abgesicherten Gesamtbild zu vereinen. Hierzu wäre es wünschenswert, Konventionen darüber zu entwickeln, wie Validierungsvergleiche durchgeführt und dokumentiert werden sollen. Diese sollten folgende Elemente enthalten.

(a) *Ein (partielles) Datengenerierungsmodell.* Bei der Untersuchung von Erinnerungsfehlern unter realistischen Bedingungen ist eine klare Vorstellung über die an der Datengenerierung beteiligten Schritte und die dort ablaufenden kognitiven Prozesse notwendig, um Erwartungen über das Auftreten von Fehlern abzuleiten, Abweichungen zu interpretieren und Verbesserungsmaßnahmen zu konzipieren. Beim Vergleich von Befragungs- mit Registerdaten ist auch deren Generierungsprozess zu berücksichtigen und die Vergleichbarkeit an allen relevanten Punkten (der Bezugseinheiten, der Vergleichszeiträume, der Informationen und der Datenformate und -strukturen) differenziert zu evaluieren. Besonders die vielfältigen Faktoren, die in der realen Welt mit dem Verstreichen von Zeit einhergehen, müssen differenziert berücksichtigt und ihre einzelnen und kombinierten Effekte auseinander gehalten werden.

(b) *Qualitative Einzelfallanalysen.* Bei komplexen und vielschichtigen Daten vermittelt eine qualitative Einzelfallanalyse ein anschauliches Verständnis davon, wie die Abweichungen aussehen und zu Stande kommen. Um die Häufigkeit und Bedingungen des Auftretens verschiedener Arten von Abweichungen zu ermitteln und ihre Folgen für die Information in den Daten abzuschätzen, wäre es aufschlussreich, eine größere Stichprobe von Personen, Episoden oder Sequenzabschnitten den verschiedenen Arten von Abweichungen zuzuordnen. Leider lässt sich die qualitative Betrachtung, wie ich sie hier mit Rückgriff auf Textfelder, Originalfragebögen und Tonbänder betrieben habe, nicht automatisieren.

(c) *Differenzierte und systematische Fehlerbeschreibung.* Zum ersten betrifft dies die Unterscheidung der beiden *Vergleichsrichtungen*, die ich als *Irrtum* und *Fehldiagnosen* bezeichnet habe. Dies ist vor allem dort sinnvoll, wo Ereignisse in sehr unterschiedlicher Häufigkeit auftreten und Irrtum und Fehldiagnosenrate deshalb sehr unterschiedlich ausfallen können, wie es zum Beispiel in Kapitel 5 für Arbeitslosigkeitsepisoden der Fall ist (siehe Tab. 5.3). Lebensläufe ohne Arbeitslosigkeit haben ein geringes Irrtumsrisiko; nur 6,7 Prozent enthalten in der zweiten Befragung abweichend eine oder zwei Arbeitslosigkeiten. Da aber weit über 90 Prozent aller Lebensläufe in der ersten Befragung keine Arbeitslosigkeit enthalten, bedeutet das, dass es sich dabei immerhin um 62 Personen handelt, die in der zweiten Befragung „fehldiagnostiziert“ werden – und damit zu knapp 37 Prozent Fehldiagnosen in der Gruppe derer mit mindestens einer Arbeitslosigkeit im Lebenslauf führen.

Zum zweiten betrifft dies die Fehleraspekte *Verbreitung, Größe, Richtung* und/oder *Abhängigkeit* vom tatsächlichen Sachverhalt, die weitgehend unabhängig voneinander auftreten und von verschiedenen Faktoren in unterschiedlichem Maß und auf unterschiedlichen Wegen beeinflusst werden. Besonders der Richtung des Fehlers sollte Aufmerksamkeit geschenkt werden, denn da Fehler in alle möglichen Richtungen auftreten können, bestimmt das relative Überwiegen einer der Richtungen den Fehler in Stichprobenkennwerten (siehe nächster Punkt).

Als drittes betrifft dies die Unterscheidung von Abweichungen *individueller* Angaben und deren Verteilung innerhalb einer Stichprobe einerseits und von *deskriptiven Kennwerten auf Stichprobenebene* andererseits. Tendenziell fallen Vergleiche auf Stichprobenebene günstiger aus als auf der Individualebene, da sich Abweichungen in entgegengesetzte Richtungen gegeneinander aufheben können.

(d) *Vergleichende Modellrechnungen.* Als Ergänzung zur Fehlerbeschreibung kann es aufschlussreich sein, regressionsanalytische Modellschätzungen auf der Grundlage der zu vergleichenden Datensätze getrennt zu berechnen und die Konsistenz der Modellparameter zu prüfen. Die Aussagekraft eines solchen Vergleichs muss allerdings realistisch betrachtet werden. Zahllose Informationen können als abhängige oder unabhängige Variablen aufgenommen oder verworfen werden, und es bleibt ungeklärt, ob Unterschiede oder Übereinstimmungen bestehen bleiben, wenn weitere Personen oder Variablen aufgenommen oder die gegenwärtigen Variablen transformiert werden. Auch im Modellvergleich in Abschnitt 5.3.7 fallen die Modellparameter deutlich diskrepanter aus, wenn die zensierten Fälle berücksichtigt werden.

(e) *Validität und Angemessenheit.* Es ist sinnvoll, die Validität bei der Erfassung sozialwissenschaftlicher Tatbestände als Angemessenheit zu betrachten, mit der ein individuelles Leben in den Kategorien der zu erfassenden sozialen Tatbestände beschrieben wird (Rohwer & Pötter, 2002). In Reliabilitätsstudien kann es dabei durchaus sinnvoll sein, die früheren Angaben im Hinblick auf den Gedächtnisfehler als den überlegenen Standard zu betrachten. Dies ist aber nicht immer der Fall und sollte deshalb zumindest begründet werden. Auch beim Vergleich von Befragungs- mit Registerdaten können die Registerdaten nicht ungeprüft als uneingeschränkt angemessener Zugang zu dem interessierenden sozialwissenschaftlichen Tatbestand erklärt werden.

In jedem Fall ist absolut unverzichtbar, die Abweichungen unter genauer Kenntnis und Berücksichtigung beider Datenkonstruktionsprozesse sowie der konzeptuellen und theoretischen Hintergründe *beider* Datensätze zu ermitteln und zu interpretieren. Wenn die Datenkonstruktionsprozesse und die zu erhebenden Konzepte in zwei Datensätzen in wesentlichen Punkten voneinander abweichen, ist es sinnvoll, den Vergleich ganz oder teilweise nicht ein-

seitig als Validierung des einen am anderen Datensatz zu führen, sondern die Abbildung ein und desselben Konzepts in den beiden Datensätzen unter den Gesichtspunkten einer *differenziellen* oder *komplementären Validität im Hinblick auf den zu erfassenden Tatbestand* zu betrachten.

#### *Schlusswort*

Die Frage nach der Qualität von Befragungsdaten stellt sich besonders in Zeiten, in denen das Vertrauen in Befragungen in der allgemeinen Bevölkerung schwindet (und sich in nachlassender Teilnahmebereitschaft und im Verkaufserfolg von Büchern mit Titeln wie „So lügt man mit Statistik“ [Krämer, 2000] niederschlägt) und zum Teil auch unter Sozialwissenschaftlern nachlässt (Brückner, 1995). Deshalb halte ich es für wichtig, sowohl die Möglichkeiten als auch die Beschränkungen von Daten gelassen und konstruktiv im Rahmen einer prozessbegleitenden Qualitätskontrolle zu thematisieren. In diesem Sinne möchte ich auch meine Untersuchung nicht als eine Kritik oder Verteidigung retrospektiver Daten betrachten. Mein Anliegen war, eine detaillierte Darstellung und Diskussion ihrer spezifischen Stärken und Schwächen anzufertigen, wie sie generell für *alle* Daten – aus Befragungen oder anderen Quellen – existieren und zugänglich sein sollten.

Solche Darstellungen sind besonders wichtig, wenn mit Daten aus verschiedenen Quellen gearbeitet wird, etwa bei einander ergänzenden Analysen zum selben Thema oder bei der Zusammenführung von Datensätzen auf Individualebene. Als Bestandteil der politischen Initiativen zur Verbesserung der informationellen Infrastruktur, die sich zum Beispiel in der Einsetzung einer Kommission aus Wissenschaftlern und Statistikern durch das Bundesministerium für Bildung und Forschung im Jahre 1999 oder der Förderung einer *Konferenz für Sozial- und Wirtschaftsdaten* im Januar 2003 mit Mitteln des Bundes zeigen, kann es nur nützlich sein, wenn mit den komplexen Datensätzen aus der amtlichen Statistik, aus Verwaltungsregistern und aus Befragungen auch deren Produktionsprozesse zugänglich und transparent gemacht werden.





## Literatur

- Abbott, Andrew, & Tsay, Angela. (2000). Sequence analysis an optimal matching methods in sociology. *Sociological Methods and Research*, 29 (1), 3–33.
- Atkinson, Richard C., & Shiffrin, Richard M. (1968). Human memory: A proposed system and its control processes. In Kenneth W. Spence (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 2, pp. 89–195). New York: Academic Press.
- Auriat, Nadia. (1992). Who forgets? An analysis of memory effects in a retrospective survey of migration history. *European Journal of Population*, 7, 311–342.
- Auriat, Nadia. (1993). „My wife knows best“ – A comparison of event dating accuracy between the wife, the husband, the couple and the Belgium population register. *Public Opinion Quarterly*, 57, 165–190.
- Auriat, Nadia. (1996). *Les défaillances de la mémoire humaine: Aspects cognitifs des enquêtes rétrospectives*. Paris: Presses Universitaires de France.
- Baddeley, Alan, Lewis, V. J., & Nemo-Smith, J. (1978). When did you last ...? In Donald Camburn (Ed.), *Practical aspects of memory* (pp. 77–83). London: Academic Press.
- Bahrnick, Harry P. (1998). Loss and distortion of memory content. In Charles P. Thompson, Douglas J. Herrmann, Darryl Bruce, J. Don Read, David G. Payne, & Michael P. Toglia (Eds.), *Autobiographical memory: Theoretical and applied perspectives* (pp. 69–78). Mahwah, NJ: Erlbaum.
- Bahrnick, Harry P., Bahrnick, Peter C., & Wittlinger, Roy P. (1975). Fifty years of memories for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General*, 113, 1–29.
- Bahrnick, Harry P., Hall, Lynda K., & Berger, Stephanie A. (1996). Accuracy and distortion in memory for high school grades. *Psychological Science*, 7 (5), 265–271.
- Balan, Jorge, Browning, Harley L., Jelin, Elizabeth, & Litzler, Lee. (1969). A computerized approach to the processing and analysis of life histories obtained in sample surveys. *Behavioral Science*, 14 (2), 105–120.
- Baltes, Paul B., & Baltes, Margret M. (1990). Psychological perspectives on successful aging: The model of selective optimization with compensation. In Paul B. Baltes & Margret M. Baltes (Eds.), *Successful aging: Perspectives from the behavioral sciences* (pp. 1–34). New York: Cambridge University Press.
- Baltes, Paul B., Lindenberger, Ulman L., & Staudinger, Ursula M. (1998). Life-span theory in developmental psychology. In Richard M. Lerner (Ed.), *Handbook of child psychology: Vol. 1. Theoretical models of human development* (pp. 1029–1143). New York: Wiley.
- Barclay, Craig R. (1986). Schematization of autobiographical memory. In David C. Rubin (Ed.), *Autobiographical memory* (pp. 82–99). Cambridge, UK: Cambridge University Press.
- Barsalou, Lawrence W. (1988). The content and organization of autobiographical memories. In Ulric Neisser & Eugene Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 193–243). New York: Cambridge University Press.
- Bartlett, Frederic Charles. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.
- Becker, Rolf. (1993). *Staatsexpansion und Karrierechancen. Berufsverläufe im öffentlichen Dienst und in der Privatwirtschaft*. Frankfurt a.M.: Campus.
- Becker, Rolf. (2001). Reliabilität von retrospektiven Berufsverlaufsdaten – ein Vergleich zwischen der Privatwirtschaft und dem öffentlichen Dienst anhand von Paneldaten. *ZUMA-Nachrichten*, 49, 29–56.
- Belli, Robert F. (1998). The structure of autobiographical memory and the event history calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory*, 6 (4), 383–406.
- Belli, Robert F., Shay, William, & Stafford, Frank. (1999). *Computerized event history calendar methods: A demonstration of features, functions and flexibility*. Paper presented at the annual meeting of the American Association for Public Opinion Research, St. Pete Beach, Florida.
- Belli, Robert F., Shay, William, & Stafford, Frank. (2001). Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly*, 65 (1), 45–74.

- Bender, Stefan, Hilzendege, Jürgen, Rohwer, Götz, & Rudolph, Helmut. (1996). *Die LAB-Beschäftigtenstichprobe 1975–1990*. Nürnberg: Institut für Arbeitsmarkt- und Berufsforschung (Beiträge zur Arbeitsmarkt- und Berufsforschung).
- Berger, Horst. (1995). Erwerbssituation ostdeutscher Haushalte. In Hans-Jürgen Andreß (Hrsg.), *Fünf Jahre danach: Zur Entwicklung von Arbeitsmarkt und Sozialstruktur im vereinten Deutschland* (S. 43–65). Berlin: de Gruyter.
- Berger, Peter A., & Sopp, Peter. (1992). Bewegtere Zeiten? Zur Differenzierung von Erwerbsverlaufsmustern in Westdeutschland. *Zeitschrift für Soziologie*, 21 (3), 166–185.
- Bick, Wolfgang, Mann, Reinhard, & Müller, Paul J. (Hrsg.). (1984). *Sozialforschung und Verwaltungsdaten*. Stuttgart: Klett-Cotta (Historisch-Sozialwissenschaftliche Forschung).
- Bick, Wolfgang, & Müller, Paul J. (1980). The nature of process-produced data: Towards a social-scientific source criticism. In Jerome M. Clubb & Erwin K. Scheuch (Eds.), *The use of historical and process-produced data* (pp. 369–413). Stuttgart: Klett-Cotta.
- Bick, Wolfgang, & Müller, Paul J. (1984). Sozialwissenschaftliche Datenkunde für prozessproduzierte Daten: Entstehungsbedingungen und Indikatorenqualität. In Wolfgang Bick, Reinhard Mann, & Paul J. Müller (Hrsg.), *Sozialforschung und Verwaltungsdaten* (S. 123–159). Stuttgart: Klett-Cotta (Historisch-Sozialwissenschaftliche Forschung).
- Bird, Katherine, Born, Claudia, & Erzberger, Christian. (2000). *Ein Bild des eigenen Lebens zeichnen: Der Kalender als Visualisierungsinstrument zur Erfassung individueller Lebensverläufe*. Bremen: Universität Bremen (Sonderforschungsbereich 186, Arbeitspapier 21).
- Blair, Edward A., & Burton, Scott. (1987). Cognitive processes used by survey respondents to answer behavioral frequency questions. *Journal of Consumer Research*, 14, 208–288.
- Blossfeld, Hans-Peter, Hamerle, Alfred, & Mayer, Karl Ulrich. (1991). Event-history models in social mobility research. In David Magnusson, Lars R. Bergman, Georg Rudinger, & Bertil Törrestad (Eds.), *Problems and methods in longitudinal research* (pp. 212–235). Cambridge, UK: Cambridge University Press.
- Blossfeld, Hans-Peter, & Rohwer, Götz. (1995). *Techniques of event history modeling: New approaches to causal analysis*. Mahwah, NJ: Erlbaum.
- Bluck, Susan. (2001). Autobiographical memories: A building block of life narratives. In Gary M. Kenyon, Brian de Vries, & Phillip Clark (Eds.), *Narrative gerontology: Theory, research and practice* (pp. 61–87). New York: Springer.
- Bluck, Susan, & Habermas, Tilmann. (2000). The Life Story Schema. *Motivation & Emotion*, 25 (2), 121–147.
- Bluck, Susan, Levine, Linda J., & Laulhere, Tracy M. (1999). Autobiographical remembering and hypernesia: A comparison of older and younger adults. *Psychology and Aging*, 14 (4), 671–682.
- Bortz, Jürgen. (1993). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Bortz, Jürgen, & Döring, Nicola. (1995). *Forschungsmethoden und Evaluation für Sozialwissenschaftler*. Berlin: Springer.
- Bower, G. H., & Reitman, J. C. (1972). Mnemonic elaboration in multilist learning. *Journal of Verbal Learning and Verbal Behavior*, 11, 478–489.
- Bradburn, Norman M., Rips, Lance J., & Shevell, Steven K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, 236, 157–161.
- Brewer, William F. (1986). What is autobiographical memory? In David C. Rubin (Ed.), *Autobiographical memory* (pp. 25–49). Cambridge, UK: Cambridge University Press.
- Brewer, William F. (1988). Memory for randomly sampled autobiographical events. In Ulric Neisser & Eugene Winograd (Eds.), *Remembering reconsidered: Ecological and traditional approaches to the study of memory* (pp. 21–90). New York: Cambridge University Press.
- Brewer, William F. (1996). What is recollective memory? In David C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 19–65). Cambridge, UK: Cambridge University Press.
- Brewer, William F., & McNamara, G. V. (1984). The nature and function of schemas. In Robert S. Wyer & Thomas K. Srull (Eds.), *Handbook of social cognition* (p. 1). Hillsdale, NJ: Erlbaum.
- Brown, Norman R. (1990). Organization of public events in long-term memory. *Journal of Experimental Psychology: General*, 119, 297–314.
- Brown, Norman R., Rips, Lance J., & Shevell, Steven L. (1985). The subjective dates of natural events in very long-term memory. *Cognitive Psychology*, 17, 139–177.
- Brown, Norman R., & Schopflocher, Donald. (1998). Event clusters: An organization of personal events in autobiographical memory. *Psychological Science*, 9 (6), 470–475.

- Brown, Norman R., Shevell, Steven K., & Rips, Lance J. (1986). Public memories and their personal context. In David C. Rubin (Ed.), *Autobiographical memory* (pp. 137–158). Cambridge, UK: Cambridge University Press.
- Brown, Roger, & Kulik, James. (1977). Flashbulb memories. *Cognition*, 5, 73–99.
- Brückner, Erika. (1993). *Lebensverläufe und gesellschaftlicher Wandel. Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1919–1921*. Berlin: Max-Planck-Institut für Bildungsforschung (Materialien aus der Bildungsforschung 44).
- Brückner, Hannah. (1995). *Surveys don't lie, people do? An analysis of data quality in a retrospective life course study*. Berlin: Max Planck Institute for Human Development.
- Brückner, Hannah, & Mayer, Karl Ulrich. (1995). Lebensverläufe und gesellschaftlicher Wandel. Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1954–1956 und 1959–1961. Berlin: Max-Planck-Institut für Bildungsforschung (Materialien aus der Bildungsforschung 48).
- Burke, Deborah M., & Light, Leah L. (1981). Memory and aging: The role of retrieval processes. *Psychological Bulletin*, 90, 513–546.
- Burt, Christopher D. B. (1992). Retrieval characteristics of autobiographical memories: Event and date information. *Applied Cognitive Psychology*, 6, 389–404.
- Burt, Christopher D. B., Kemp, Simon, & Conway, Martin. (2001). What happens if you retest autobiographical memory 10 years on? *Memory & Cognition*, 29 (1), 127–136.
- Cannell, Charles F., Fisher, Gordon, & Bakker, Thomas. (1965). *Reporting of hospitalization in the Health Interview Survey*. Washington, DC: National Center for Health Statistics.
- Cash, William S., & Moss, Abigail J. (1972). *Optimum recall period for reporting persons injured in motor vehicle accidents*. Washington, DC: National Center for Health Statistics.
- Caspi, Avshalom, Moffitt, Terry E., Thornton, Arland, Freedman, Deborah, Amell, James W., Harrington, Honalee L., Smeijers, Judith, & Silva, Phil A. (1996). The life history calendar: A research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research*, 6 (2), 101–114.
- Cattell, Raymond. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton Mifflin.
- Ceci, Stephen J., Ross, David, & Toglia, Michael. (1987). Age differences in suggestibility: Psychological implications. *Journal of Experimental Psychology: General*, 118, 38–49.
- Christianson, Sven-Ake, & Loftus, Elisabeth F. (1991). Remembering emotional events: The fate of detailed information. *Cognition and Emotion*, 5, 81–108.
- Cohen, Gillian. (1998). The effects of aging on autobiographical memory. In Charles P. Thompson, Douglas J. Herrmann, Darryl Bruce, J. Don Read, David G. Payne, & Michael P. Toglia (Eds.), *Autobiographical memory: Theoretical and applied perspectives* (pp. 105–123). Mahwah, NJ: Erlbaum.
- Cohen, Gillian, Conway, Martin A., & Maylor, E. (1994). Flashbulb memories in older adults. *Psychology and Aging*, 9, 454–463.
- Coleman, Peter G. (1991). Ageing and life history: The meaning of reminiscence in later life. In Shirley Dex (Ed.), *Life and work history analyses: Qualitative and quantitative developments* (pp. 120–143). London: Routledge.
- Conrad, Frederick G., & Schober, Michael F. (2000). Clarifying question meaning in a standardized survey. *Public Opinion Quarterly*, 64, 1–28.
- Conway, Martin A. (1996). Autobiographical knowledge and autobiographical memory. In David C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 67–93). Cambridge, UK: Cambridge University Press.
- Conway, Martin A., & Bekerian, Debra A. (1987). Organization in autobiographical memory. *Memory & Cognition*, 15, 119–135.
- Conway, Martin A., & Pleydell-Pearce, Christopher W. (2000). The construction of autobiographical memories in the self-memory system. *Psychological Review*, 107 (2), 261–288.
- Courgeau, Daniel. (1991). Analyse des données biographiques erronées. *Population*, 1, 89–104.
- Crovitz, Herbert F., & Schiffman, H. (1974). Frequency of episodic memories as a function of their age. *Bulletin of the Psychonomic Society*, 4, 517–518.
- Dex, Shirley, & McCulloch, Andrew. (1998). The reliability of retrospective unemployment history data. *Work, Employment & Society*, 12, 3, 497–509.
- Diekmann, Andreas. (1995). *Empirische Sozialforschung. Grundlagen, Methoden, Anwendungen*. Reinbek: Rowohlt.

- Diewald, Martin. (1999). *Entwertungen, Umwertungen, Aufwertungen – Ostdeutsche Erwerbsverläufe zwischen Kontinuität und Wandel*. Berlin: Freie Universität.
- Dijkstra, Wil, & Taris, Toon. (1995). Measuring the agreement between sequences. *Sociological methods and research*, 24 (2), 214–231.
- Ebbinghaus, Hermann. (1885/1964). *Memory: A contribution to experimental psychology*. New York: Dover.
- Elias, Peter. (1991). Methodological, statistical and practical issues arising from the collection and analyses of work history information by survey techniques. *Bulletin de Méthodologie Sociologique*, 31, 3–31.
- Elias, Peter. (1997). *Who forgot they were unemployed?* University of Essex: ESRC Research Centre on Micro-Social Change (Working Paper 97).
- Ericsson, Karl Anders, & Chase, William G. (1982). Exceptional memory. *American Scientist*, 70 (6), 607–615.
- Fienberg, Stephen E., & Tanur, Judith M. (1989). Combining cognitive and statistical approaches to survey design. *Science*, 243, 1017–1023.
- Fitzgerald, Joseph M. (1996). Intersecting meanings of reminiscence in adult development and aging. In David C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 360–381). Cambridge, UK: Cambridge University Press.
- Fitzgerald, Joseph M., & Lawrence, Renee. (1984). Autobiographical memory across the life span. *Journal of Gerontology*, 39, 692–698.
- Fivush, Robyn, Gray, J. T., & Fromhoff, F. A. (1987). Two-year olds talk about the past. *Cognitive Development*, 2, 393–409.
- Fivush, Robyn, Haden, C., & Adam, S. (1995). Structure and coherence of preschoolers' personal narratives over time: Implications for childhood amnesia. *Journal of Experimental Child Psychology*, 60, 32–56.
- Fowler, Floyd J., & Roman, Anthony M. (1992). *A study of approaches to survey question evaluation*. Boston: University of Massachusetts.
- Freedman, Deborah A., Thornton, Arland, Camburn, Donald, Alwin, Duane F., & Young-DeMarco, Linda. (1988). The life history calendar: A technique for collecting retrospective data. *Sociological Methodology*, 18, 37–68.
- Friedman, William J. (1993). Memory for the time of past events. *Psychological Bulletin*, 113, 44–66.
- Friedman, William J., & Pines, A. (1991). Sex differences in gender-related childhood memories. *Sex Roles*, 25, 25–32.
- Friedman, William J., & Wilkins, Arnold J. (1985). Scale effects in memory for the time of events. *Memory & Cognition*, 13, 168–175.
- Fromholt, Pia, & Larsen, Steen F. (1994). Autobiographical memory in normal aging and primary degenerative dementia (dementia of Alzheimer type). *Journal of Gerontology*, 46 (3), 85–91.
- Gallie, Duncan, & Russell, Helen. (1998). Unemployment and life satisfaction: A cross-cultural comparison. *European Journal of Sociology*, 2, 248–280.
- Gallie, Duncan, & Vogler, Carolyn. (1994). Unemployment and attitudes to work. In Duncan Gallie, Catherine Marsh, & Carolyn Vogler (Eds.), *Social change and the experience of unemployment* (pp. 116–153). Oxford: Oxford University Press.
- Gergen, Kenneth J., & Gergen, M. M. (1988). Narrative and the Self as relationship. In Lynne Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 19–56). San Diego, CA: Academic Press.
- Gibbons, Jeffrey A., & Thompson, Charles P. (2001). Using a calendar in event dating. *Applied Cognitive Psychology*, 15, 33–44.
- Goedicke, Anne. (2000). Beschäftigungsfolgen betrieblichen Wandels. Eine Studie zu Betriebswechsellern und Arbeitslosigkeit in Ostdeutschland. Berlin: Freie Universität.
- de Graaf, Nan Dirk, & Wegener, Bernd. (1989). *(Un)reliability of job career and social resource data*. Unpublished manuscript, Max Planck Institute for Human Development, Berlin.
- Groninger, L. D. (1971). Mnemonic imagery and forgetting. *Psychonomic Science*, 23 (2), 161–163.
- Groves, Robert M. (1989). *Survey errors and survey costs*. New York: Wiley.
- Groves, Robert M. (1996). How do we know what we think they think is really what they think? In Norbert Schwarz & Seymour Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research* (pp. 389–402). San Francisco: Jossey-Bass.
- Gruneberg, Michael M. (1976). The distinction between short-term memory and long-term memory. *Bulletin of the British Psychological Society*, 29, 327–333.

- Gruneberg, Michael M., Morris, Peter E., & Sykes, Robert N. (Eds.). (1978). *Practical aspects of memory*. Chichester, UK: Wiley.
- Gruneberg, Michael M., Morris, Peter E., & Sykes, Robert N. (Eds.). (1988). *Practical aspects of memory: Current research and issues: Vol. 1. Memory in everyday life*. Chichester, UK: Wiley.
- Habermas, Tilmann, & Bluck, Susan. (2000). Getting a life: The emergence of the Life Story Schema in adolescence. *Psychological Bulletin*, 126 (5), 748–769.
- Halbwachs, Maurice. (1985). *Das kollektive Gedächtnis*. Frankfurt a.M.: Fischer Taschenbuch Verlag.
- Hamilton, J. C., & Shuminsky, T. R. (1990). Self-awareness mediates the relationship between serial position and item reliability. *Journal of Personality and Social Psychology*, 59, 1301–1307.
- Harre, Elke, & Schmidt, Ines. (1995). *Wandel und Kontinuitäten in den Haltungen Ostberliner Mütter und Töchter zu Erwerbstätigkeit und Familie*. Berlin: trafo Verlag.
- Heaton, Tim B. (1990). Marital stability throughout the childrearing years. *Demography*, 27 (1), 55–63.
- Heckhausen, Jutta. (1993). Developmental expectations for the self and most other people: Age grading in three functions of social comparison. *Developmental Psychology*, 29, 539–548.
- Helmchen, Hanfried, Baltes, Margret M., Geiselmann, Bernhard, Kanowski, Siegfried, Linden, Michael, Reischies, Friedel, Wagner, Michael, Wernicke, Thomas, & Wilms, Hans-Ulrich. (1996). Psychische Erkrankungen im Alter. In Karl Ulrich Mayer & Paul B. Baltes (Hrsg.), *Die Berliner Altersstudie* (S. 185–219). Berlin: Akademie Verlag.
- Hillmert, Steffen. (2002). *Die Edition von Lebensverlaufsdaten – Einzelfallprüfungen, Korrekturentscheidungen und ihre Relevanz*. Berlin: Max-Planck-Institut für Bildungsforschung (Projekt Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland, Arbeitspapier 20).
- Hillmert, Steffen, Künster, Ralf, Spengemann, Petra, & Mayer, Karl Ulrich. (2004). *Projekt „Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland“*. Dokumentation: Teil IV. Editionsbandbuch. Berlin: Max-Planck-Institut für Bildungsforschung (Materialien aus der Bildungsforschung 78).
- Holst, Elke (2000). *Die Stille Reserve am Arbeitsmarkt: Größe – Zusammensetzung – Verhalten*. Berlin: edition sigma.
- Houtkoop-Steenstra, Hanneke. (2000). *Interaction and the standardized survey interview: The living questionnaire*. Cambridge, UK: Cambridge University Press.
- Huinink, Johannes. (1995). *Warum noch Familie? Zur Attraktivität von Partnerschaft und Elternschaft in unserer Gesellschaft*. Frankfurt a.M.: Campus.
- Hullen, Gert. (1995). Der Auszug aus dem Elternhaus im Vergleich von West- und Ostdeutschland. Ergebnisse des Family and Fertility Surveys (FFS) 1992. *Zeitschrift für Bevölkerungswissenschaften*, 20 (2), 141–158.
- Hullen, Gert. (1998). *Lebensverläufe in West- und Ostdeutschland. Längsschnittanalysen des deutschen Family and Fertility Survey*. Opladen: Leske + Budrich.
- Huttenlocher, Jannellen, Hedges, Larry V., & Bradburn, Norman M. (1990). Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology, Learning, Memory and Cognition*, 16, 196–213.
- Huttenlocher, Jannellen, Hedges, Larry V., & Prohaska, Vincent. (1988). Hierarchical organization in ordered domains: Estimating the dates of events. *Psychological Review*, 95 (4), 471–484.
- infas. (1999). *Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland. Methodenbericht zur Hauptstudie*. Bonn: Institut für Angewandte Sozialwissenschaft.
- Jabine, Thomas B., Straf, Miron L., Tanur, Judith M., & Tourangeau, Roger (Eds.). (1984). *Cognitive aspects of survey methodology: Building a bridge between disciplines*. Washington, DC: National Academy Press.
- Jacobs, Sheila C. (2002). Reliability and recall of unemployment events using retrospective data. *Work, Employment and Society*, 16 (3), 537–548.
- Jahoda, Marie. (1982). *Employment and unemployment: A social psychological analysis*. Cambridge, UK: Cambridge University Press.
- James, William. (1904). *The principles of psychology*. New York: Holt.
- Jobe, Jared B., Keller, Donald M., & Smith, Albert F. (1996). Cognitive techniques in interviewing older people. In Norbert Schwarz & Seymour Sudman (Eds.), *Thinking about answers: The application of cognitive processes to survey methodology* (pp. 197–219). San Francisco: Jossey-Bass.
- Jobe, Jared B., & Mingay, David J. (1991). Cognition and survey measurement: History and overview. *Applied Cognitive Psychology*, 5, 175–192.

- Jobe, Jared B., Tourangeau, Roger, & Smith, A. F. (1993). Contributions of survey research to the understanding of memory. *Applied Cognitive Psychology*, 7, 567–584.
- Johnson, Robert A., Gerstein, Dean R., & Rasinski, Kenneth A. (1998). Adjusting survey estimates for response bias: An application to trends in alcohol and marijuana use. *Public Opinion Quarterly*, 62 (3), 354–377.
- Kalicki, Bernhard. (1996). *Lebensverläufe und Selbstbilder: Die Normalbiographie als psychisches Regulativ*. Opladen: Leske + Budrich.
- Klein, Thomas, & Fischer-Kerli, David. (2000). Die Zuverlässigkeit retrospektiv erhobener Lebensverlaufsdaten. Analysen zur Partnerschaftsbiografie des Familiensurveys. *Zeitschrift für Soziologie*, 29 (4), 294–312.
- Kliegl, Reiner, Smith, Jacqui, Heckhausen, Jutta, & Baltes, Paul B. (1987). Mnemonic training for the acquisition of skilled digit memory. *Cognition & Instruction*, 4 (4), 203–223.
- Klijzing, Eric, & Prophet, Helen. (1998). *On the quality of FFS event history data*. Geneva: UN Economic Commission for Europe, Population Activities Unit.
- Knowles, E. S. (1988). Item context effects on personality scales: Measuring changes in the measure. *Journal of Personality and Social Psychology*, 55, 312–330.
- Kohli, Martin. (1985). Die Institutionalisierung des Lebenslaufs. Historische Befunde und theoretische Argumente. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 37, 1–29.
- Krämer, Walter. (2000). *So lügt man mit Statistik*. München: Piper.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, Amos. (1971). *Foundations of measurement: Vol. 1. Additive and polynomial representations*. New York: Wiley.
- Kruskal, William. (1991). Introduction. In Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz & Seymour Sudman (Eds.), *Measurement error in surveys* (pp. xxiii–xxxiii). New York: Wiley.
- Kühnemund, Harald. (1990). „Wie war das doch gleich ...“ *Zur Problematik retrospektiver Befragungen*. Berlin: Freie Universität.
- Lampinen, James M., Faries, Jeremiah M., Neuschatz, Jeffrey S., & Togli, Michael P. (2000). Recollection of things schematic: The influence of scripts on recollective experience. *Applied Cognitive Psychology*, 14, 543–554.
- Larsen, Steen F., & Thompson, Charles P. (1995). Reconstructive memory in the dating of personal and public events. *Memory & Cognition*, 23 (6), 780–790.
- Larsen, Steen F., Thompson, Charles P., & Hansen, Tia. (1995). Time in autobiographical memory. In David C. Rubin (Ed.), *Remembering our past: Studies in autobiographical memory* (pp. 129–156). Cambridge, UK: Cambridge University Press.
- Lauterbach, Wolfgang. (1994). *Berufsverläufe von Frauen. Erwerbstätigkeit, Unterbrechung und Wiedereintritt*. Frankfurt a.M.: Campus.
- Lauterbach, Wolfgang, & Lüscher, Kurt. (1999). *Wer sind die Spätauszieger? Oder: Herkunftsfamilie, Wohnumfeld und die Gründung eines eigenen Haushaltes*. Konstanz: Universität, Sozialwissenschaftliche Fakultät.
- Lessler, Judith T., & Forsyth, Barbara H. (1996). A coding system for appraising questionnaires. In Seymour Sudman, Norman M. Bradburn, & Norbert Schwarz (Eds.), *Thinking about answers: The application of cognitive processes to survey methodology* (pp. 259–291). San Francisco: Jossey-Bass.
- Lieury, Alain, Aiello, Bruna, Lepreux, Dominique, & Mellet, Monique. (1980). Le rôle de repères dans la récupération et la datation des souvenirs. *Année Psychologique*, 80, 149–167.
- Lieury, Alain, Richer, E., & Weeger, I. (1978). Les événements privés et publics dans la datation des souvenirs anciens. *Bulletin de Psychologie*, 32 (338), 41–48.
- Linton, Marigold. (1975). Memory for real world events. In Donald A. Norman & David E. Rumelhart (Eds.), *Explorations in cognition* (pp. 376–404). San Francisco: Freeman.
- Linton, Marigold. (1982). Transformations of memory in everyday life. In Ulric Neisser (Ed.), *Memory observed: Remembering in natural contexts* (pp. 77–91). San Francisco: Freeman.
- Linton, Marigold. (1986). Ways of searching and the contents of memory. In David C. Rubin (Ed.), *Autobiographical memory* (pp. 50–67). Cambridge, UK: Cambridge University Press.
- Loftus, Elizabeth F., & Fathi, D. C. (1985). Retrieving multiple autobiographical memories. *Social Cognition*, 3 (3), 280–295.
- Loftus, Elizabeth F., Fienberg, Stephen E., & Tanur, Judith M. (1985). Cognitive Psychology meets the National Survey. *American Psychologist*, 40 (2), 157–180.

- Loftus, Elizabeth F., & Marburger, Wesley. (1983). „Since the eruption of Mt. St. Helens, has anyone beaten you up?“ Improving the accuracy of retrospective reports with landmark events. *Memory & Cognition*, 11 (2), 114–120.
- Lord, Frederic M., & Novick, Melvin R. (1968). *Statistical theory of mental test scores*. Reading, MA: Addison-Wesley.
- Maas, Ineke, Borchelt, M., & Mayer, Karl Ulrich. (1996). Kohortenschicksale der Berliner Alten. In Karl Ulrich Mayer & Paul B. Baltes (Hrsg.), *Die Berliner Altersstudie* (S. 132–134). Berlin: Akademie Verlag.
- Marcoux, R. (1997). *Amélioration qualitative de nos outils quantitatifs. Deux exemples: La datation des événements et le statut d'occupation résidentiel*. Conférence „L'apport des collectes biographiques pour la connaissance de la mobilité“, Paris.
- Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology*, 35, 63–78.
- Marquis, K. (1978). *Record check validity of survey responses: A reassessment of bias in reports of hospitalizations*. Chicago: Rand McNally.
- Mathiowetz, Nancy A., & Duncan, Greg J. (1988). „Out of work, out of mind“: Response errors in retrospective reports of unemployment. *Journal of Business and Economic Statistics*, 6, 221–229.
- Mayer, Karl Ulrich. (1989). Das Forschungsprojekt „Lebensverläufe und Wohlfahrtsentwicklung“. In Karl Ulrich Mayer & Erika Brückner, *Lebensverläufe und Wohlfahrtsentwicklung. Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1929–1931, 1939–1941, 1949–1951* (Teil I, S. 1–21). Berlin: Max-Planck-Institut für Bildungsforschung (Materialien aus der Bildungsforschung 35).
- Mayer, Karl Ulrich. (1998). Lebensverlauf. In Bernhard Schäfers & Wolfgang Zapf (Hrsg.), *Handwörterbuch zur Gesellschaft Deutschlands* (S. 438–451). Opladen: Leske + Budrich.
- Mayer, Karl Ulrich. (2001). Soziale Mobilität und Erwerbsverläufe in der Transformation Ostdeutschlands. In Wolfgang Schluchter & Peter E. Quint (Hrsg.), *Der Vereinigungsschock* (S. 336–365). Weilerswist: Velbrück.
- Mayer, Karl Ulrich, Baltes, Paul B., Baltes, Margret M., Borchelt, Markus, Delius, Julia, Helmchen, Hanfried, Linden, Michael, Smith, Jacqui, Staudinger, Ursula M., Steinhagen-Thiessen, Elisabeth, & Wagner, Michael. (1996). Wissen über das Alter(n). Eine Zwischenbilanz der Berliner Altersstudie. In Karl Ulrich Mayer & Paul B. Baltes (Hrsg.), *Die Berliner Altersstudie* (S. 599–634). Berlin: Akademie Verlag.
- Mayer, Karl Ulrich, & Brückner, Erika. (1989). *Lebensverläufe und Wohlfahrtsentwicklung. Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1929–1931, 1939–1941, 1949–1951*. Berlin: Max-Planck-Institut für Bildungsforschung (Materialien aus der Bildungsforschung 35).
- Mayer, Karl Ulrich, & Huinink, Johannes. (1990). Alters-, Perioden- und Kohorteneffekte in der Analyse von Lebensverläufen oder: Lexis ade? In Karl Ulrich Mayer (Hrsg.), *Lebensverläufe und sozialer Wandel* (S. 442–459). Opladen: Westdeutscher Verlag (Kölner Zeitschrift für Soziologie und Sozialpsychologie, Sonderheft 30).
- Mayer, Karl Ulrich, & Wagner, Michael. (1986). Der Auszug von Kindern aus dem elterlichen Haushalt – ein Erklärungsmodell für die Geburtsjahrgänge 1929–31, 1939–40 und 1949–50. In Klaus F. Zimmermann (Hrsg.), *Demographische Probleme der Haushaltsökonomie* (S. 2–42). Bochum: Studienverlag Dr. N. Brockmeyer.
- McGinnity, Frances. (2001). *Who benefits? A comparison of welfare and outcomes for the unemployed in Britain and Germany*. Oxford: Nuffield College, Faculty of Social Science Studies.
- McGovern, Paul G., Lurie, Nicole, Margolis, Karen L., & Slater, Jonathan S. (1998). Accuracy of self-report of mammography and pap smear in a low-income urban population. *American Journal of Preventive Medicine*, 14 (3), 201–208.
- Means, Barbara, & Loftus, Elizabeth F. (1991). When personal history repeats itself: Decomposing memories for recurring events. *Applied Cognitive Psychology*, 5, 297–318.
- Means, Barbara, Nigam, Arti, Zarrow, Marlene, Loftus, Elizabeth F., & Donaldson, M. (1989). *Autobiographical memory for health-related events*. Hyattsville, MD: U.S. Department of Health and Human Services.
- Menon, Geeta, Raghurib, Priya, & Schwarz, Norbert. (1995). Behavioral frequency judgments: An accessibility-diagnostics framework. *Journal of Consumer Research*, 22, 212–272.
- Middendorf, Elke. (2000). Panta rhei oder der mentale Fluß von Tatsachen: Zur Reliabilität retrospektiv erhobener biographischer Ereignisse. *ZA-Nachrichten*, 46, 58–71.
- Middleton, D., & Edwards, D. (Eds.). (1990). *Collective remembering*. London: Sage.



- Morgan, S. Philip, & Rindfuss, Ronald R. (1985). Marital disruption: Structural and temporal dimensions. *American Journal of Sociology*, 94 (1), 110–129.
- Mutz, Gerd. (1996). Das Problem der Versprachlichung von Arbeitslosigkeit in West- und Ostdeutschland. *BIOS – Zeitschrift für Biographieforschung*, 9 (1), 93–113.
- Neisser, Ulric. (1986). Nested structure in autobiographical memory. In David C. Rubin (Ed.), *Autobiographical memory* (pp. 71–80). Cambridge, UK: Cambridge University Press.
- Neisser, Ulric, & Fivush, Robyn (Eds.). (1994). *The remembering self: Construction and accuracy in the self-narrative*. Cambridge, UK: Cambridge University Press.
- Neisser, Ulric, & Harsch, Nicole. (1992). Phantom flashbulbs: False recollections of hearing the news about Challenger. In Eugene Winograd & Ulric Neisser (Eds.), *Affect and accuracy in recall: Studies of flashbulb memories* (pp. 9–31). New York: Cambridge University Press.
- Neter, John, Maynes, E. S., & Ramanathan, R. (1965). The effect of mismatching on the measurement of response errors. *Journal of the American Statistical Association*, 60 (312), 1005–1027.
- Neter, John, & Wakesberg, J. (1964). A study of response errors in expenditures data from household interviews. *Journal of the American Statistical Association*, 59, 18–55.
- Nickel, Hildegard Maria. (1994). Mit dem DDR-Gleichstellungsvorsprung in die bundesdeutsche Modernisierungsfalle? Deutschlands Frauen nach der Wende. *Berliner Debatte Initial*, 4, 3–14.
- Norwood, Janet L., & Tanur, Judith M. (1994). Measuring unemployment in the nineties. *Public Opinion Quarterly*, 58, 277–294.
- Ott, Norburga. (1991). Die Wirkung politischer Maßnahmen auf die Familienbildung aus ökonomischer und verhandlungstheoretischer Sicht. In Karl Ulrich Mayer, Jutta Allmendinger, & Johannes Huinink (Hrsg.), *Vom Regen in die Traufe: Frauen zwischen Beruf und Familie* (S. 385–407). Frankfurt a.M.: Campus.
- Papastefanou, Christiane. (1997). *Auszug aus dem Elternhaus: Aufbruch und Ablösung im Erleben von Eltern und Kindern*. Weinheim: Juventa.
- Papastefanou, Georgios. (1980). *Zur Güte von retrospektiven Daten – eine Anwendung gedächtnispsychologischer Theorie und Ergebnisse einer Nachbefragung*. Unveröff. Manuskript, Max-Planck-Institut für Bildungsforschung, Berlin.
- Paull, Gillian. (2002). *Biases in the reporting of labour market dynamics*. London: The Institute for Fiscal Studies, Centre for Economic Performance.
- Pierret, Charles R. (2001). Event history data and survey recall. *The Journal of Human Resources*, 36 (3), 439–466.
- Poterba, James M., & Summers, Lawrence H. (1984). Reporting errors and labor market dynamics. *Econometrica*, 54 (6), 1319–1338.
- Poulain, Michel, Riandey, Benoit, & Firdion, Jean-Marie. (1991). Enquête biographique et registre belge de population: Une confrontation des données. *Population*, 1, 89–104.
- Rabbitt, Patrick M. A., & Winthorpe, Carol A. (1988). What would old people remember? The Galton paradigm reconsidered. In Michael M. Gruneberg, Peter E. Morris, & Robert N. Sykes (Eds.), *Practical aspects of memory: Current research and issues: Vol. 1. Memory in everyday life* (pp. 301–307). Chichester, UK: Wiley.
- Reischies, Friedel M., & Lindenberger, Ulman L. (1996). Grenzen und Potentiale der kognitiven Leistungsfähigkeit im hohen Alter. In Karl Ulrich Mayer & Paul B. Baltes (Hrsg.), *Die Berliner Altersstudie* (S. 351–377). Berlin: Akademie Verlag.
- Reiser, Brian J., Black, John B., & Abelson, Robert P. (1985). Knowledge structure in the organization and retrieval of autobiographical memories. *Cognitive Psychology*, 17, 89–137.
- Reiser, Brian J., Black, John B., & Kalamarides, Peter. (1986). Strategic memory search processes. In David C. Rubin (Ed.), *Autobiographical memory* (pp. 100–121). Cambridge, UK: Cambridge University Press.
- Robinson, John A. (1992). First experience memories: Contexts and function in personal histories. In M. Conway, David C. Rubin, H. Spinnler, & W. A. Wagenaar (Eds.), *Theoretical perspectives on autobiographical memory* (pp. 223–239). Dordrecht: Kluwer.
- Robinson, John A., & Taylor, Leslie R. (1998). Autobiographical memory and self-narratives: A tale of two stories. In Charles P. Thompson, Douglas J. Herrmann, Darryl Bruce, J. Don Read, David G. Payne, & Michael P. Toglia (Eds.), *Autobiographical memory: Theoretical and applied perspectives* (pp. 125–143). Mahwah, NJ: Erlbaum.

- Rohwer, Götz. (1994). *Kontingente Lebensverläufe. Soziologische und statistische Aspekte ihrer Beschreibung und Erklärung*. Bremen: Universität Bremen.
- Rohwer, Götz, & Pötter, Ulrich. (2002). *Methoden sozialwissenschaftlicher Datenkonstruktion*. Weinheim: Juventa.
- Ross, Bruce M. (1991). *Remembering the personal past*. New York: Oxford University Press.
- Ross, Michael, & Holmberg, D. (1990). Recounting the past: Gender differences in the recall of events in the history of a close relationship. In M. P. Zanna & H. M. Olson (Eds.), *The Ontario Symposium: Vol. 6. Self-inference processes* (pp. 135–152). Hillsdale, NJ: Erlbaum.
- Rubin, David C. (1982). On the retention function for autobiographical memory. *Journal of Verbal Learning and Verbal Behavior*, 21, 21–38.
- Rubin, David C. (Ed.). (1986). *Autobiographical memory*. Cambridge, UK: Cambridge University Press.
- Rubin, David C. (Ed.). (1996). *Remembering our past: Studies in autobiographical memory*. Cambridge, UK: Cambridge University Press.
- Rubin, David C. (1998). Beginnings of a theory of autobiographical remembering. In Charles P. Thompson, Douglas J. Herrmann, Darryl Bruce, J. Don Read, David G. Payne, & Michael P. Toglia (Eds.), *Autobiographical memory: Theoretical and applied perspectives* (pp. 47–68). Mahwah, NJ: Erlbaum.
- Rubin, David C., & Baddeley, Alan. (1989). Telescoping is not time compression: A model of the dating of autobiographical events. *Memory & Cognition*, 17 (6), 653–661.
- Rubin, David C., & Kozin, Marc. (1984). Vivid memories. *Cognition*, 16, 81–95.
- Rubin, David C., & Wenzel, Amy E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, 103 (4), 734–760.
- Rubin, David C., Wetzler, Scott E., & Nebes, Robert D. (1986). Autobiographical memory across the lifespan. In David C. Rubin (Ed.), *Autobiographical memory* (pp. 202–222). Cambridge, UK: Cambridge University Press.
- Sawilowski, Shlomo S. (2000). Psychometrics vs. datametrics: Comment on Vacha-Haase's „reliability generalization“ and some EPM editorial policies. *Educational and Psychological Measurement*, 60 (2), 157–173.
- Schaeffer, Nora Cate, & Maynard, Douglas W. (1996). From paradigm to prototype and back again: Interactive aspects of cognitive processing in standardized surveys. In Norbert Schwarz & Seymour Sudman (Eds.), *Answering questions: Methodology for determining cognitive and communicative processes in survey research* (pp. 65–88). San Francisco: Jossey-Bass.
- Schömann, Klaus. (1994). *The dynamics of labor earnings over the life course: A comparative longitudinal analysis of Germany and Poland*. Berlin: Max Planck Institute for Human Development.
- Schwarz, Norbert. (1990). Assessing frequencies of mundane behavior: Contribution of Cognitive Psychology to questionnaire construction. In Clyde Hendrick & Margaret S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 98–119). Newbury Park, CA: Sage.
- Schwarz, Norbert, & Hippler, Hans J. (1991). Response alternatives: The impact of their choice and ordering. In Paul P. Biemer, Robert M. Groves, Lars E. Lyberg, Nancy A. Mathiowetz, & Seymour Sudman (Eds.), *Measurement error in surveys* (pp. 41–56). New York: Wiley.
- Schwarz, Norbert, & Sudman, Seymour (Eds.). (1996). *Answering questions: Methodology for determining cognitive and communicative processes in survey research*. San Francisco: Jossey-Bass.
- Settersten, R. A., & Mayer, Karl Ulrich. (1997). The measurement of age, age structuring, and the life course. *Annual Review of Sociology*, 26, 233–261.
- Sheppard, Leyland C., & Teasdale, John D. (1996). Depressive thinking: Changes in schematic mental models of self and world. *Psychological Medicine*, 26 (5), 1043–1051.
- Shum, M. S. (1998). The role of temporal landmarks in autobiographical memory processes. *Psychological Bulletin*, 124 (3), 423–442.
- Silberstein, Adriana. (1989). Recall effects in the US Consumer Expenditure Study. *Journal of Official Statistics*, 5 (2), 125–142.
- Singer, Jefferson A., & Salovey, Peter. (1993). *The remembered self: Emotion and memory in personality*. New York: Free Press.
- Skowronski, John J., & Thompson, Charles P. (1990). Reconstructing the dates of personal events: Gender differences in accuracy. *Applied Cognitive Psychology*, 4 (5), 371–381.
- Solga, Heike. (1996). Lebensverläufe und historischer Wandel in der ehemaligen DDR. *ZA-Information*, 38, 28–38.
- Solga, Heike. (2001). Longitudinal surveys and the study of occupational mobility: Panel and retrospective design in comparison. *Quality and Quantity*, 35 (3), 291–309.

- Squire, Larry R. (1987). *Memory and brain*. New York: Oxford University Press.
- Sudman, Seymour, & Bradburn, Norman M. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association*, 64 (344), 805–815.
- Sudman, Seymour, & Bradburn, Norman M. (1974). *Response effects in surveys: A review and synthesis*. Chicago: Aldine.
- Sudman, Seymour, & Bradburn, Norman M. (1982). *Asking questions*. San Francisco: Jossey-Bass.
- Sudman, Seymour, Bradburn, Norman M., & Schwarz, Norbert (Eds.). (1996). *Thinking about answers: The application of cognitive processes to survey methodology*. San Francisco: Jossey-Bass.
- Sudman, Seymour, Finn, A., & Lannom, L. (1984). The use of bounded recall procedures in single interviews. *Public Opinion Quarterly*, 48 (2), 520–524.
- Suessbrick, Anna, Schober, Michael F., & Conrad, Frederick G. (2000). Different respondents interpret ordinary questions quite differently. In *Proceedings of the American Statistical Association: Section on survey research methods*. Alexandria, VA: The American Statistical Association.
- Thélot, C. (1990). L'erreur de mémoire sur la situation professionnelle. *Les Collections de l'INSEE*, 175–189.
- Thompson, Bruce (1994). Guidelines for authors. *Educational and Psychological Measurement*, 54, 837–847.
- Thompson, Bruce, & Vacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement*, 60 (2), 174–195.
- Thompson, C., Skowronski, John J., & Lee, D. J. (1987). *Reconstructing the date of a personal event*. Paper presented at the Second International Conference on Practical Aspects of Memory, Swansea, Wales.
- Thompson, Charles P. (1982). Memory for unique personal events: The roommate study. *Memory & Cognition*, 48 (3), 187–201.
- Thompson, Charles P. (1998). The bounty of everyday memory. In Charles P. Thompson, Douglas J. Herrmann, Darryl Bruce, J. Don Read, David G. Payne, & Michael P. Toglia (Eds.), *Autobiographical memory: Theoretical and applied perspectives* (pp. 29–43). Mahwah, NJ: Erlbaum.
- Thompson, Charles P., Herrmann, Douglas J., Bruce, Darryl, Read, J. Don, Payne, David G., & Toglia, Michael P. (1998). *Autobiographical memory: Theoretical and applied perspectives*. Mahwah, NJ: Erlbaum.
- Thompson, Charles P., Skowronski, John J., & Lee, D. J. (1988). Telescoping in naturally occurring events. *Memory & Cognition*, 16 (3), 461–468.
- Thompson, Charles P., Skowronski, John J., Larsen, Steen F., & Betz, Andrew L. (1996). *Autobiographical memory: Remembering what and remembering when*. Mahwah, NJ: Erlbaum.
- Thornton, Arland. (1977). Children and marital stability. *Journal of Marriage and the Family*, 39 (3), 531–540.
- Tölke, Angelika. (1980). Zuverlässigkeit retrospektiver Verlaufsdaten. Qualitative Ergebnisse einer Nachbefragung. Unveröff. Manuskript, Max-Planck-Institut für Bildungsforschung, Berlin.
- Tourangeau, Roger, Rips, Lance J., & Rasinski, Kenneth A. (2000). *The psychology of survey response*. New York: Cambridge University Press.
- Trochim, William M. (2000). *The research methods knowledge base* (2nd ed.) <<http://www.socialresearch-methods.net/kb/index.htm>> (02.08.2000)
- Tulving, Endel (1972). Episodic and semantic memory. In Endel Tulving & W. Donaldson (Eds.), *Organization of memory* (pp. 381–403). New York: Academic Press.
- Tulving, Endel (1985). How many memory systems are there? *American Psychologist*, 40, 385–398.
- Tulving, Endel (1989). Memory, performance, knowledge and experience. *European Journal of Cognitive Psychology*, 1, 3–26.
- Wacker, Ali. (1976). *Arbeitslosigkeit. Soziale und psychische Voraussetzungen und Folgen*. Frankfurt a.M.: Europäische Verlagsanstalt.
- Wagenaar, Willem A. (1986). My memory: A study of autobiographical memory over six years. *Cognitive Psychology*, 18, 225–252.
- Wagenaar, Willem A. (1988). Calibration and the effects of knowledge and reconstruction in retrieval from memory. *Cognition*, 28, 277–296.
- Wagenaar, Willem A., & Groeneweg, J. (1990). The memory of concentration camp survivors. *Applied Cognitive Psychology*, 4, 77–87.
- Wagner, Michael. (1989). Korrekturen der Wohnverlaufsdaten: Ein Beispiel für die Aufbereitung von Retrospektivdaten. In Karl Ulrich Mayer & Erika Brückner, *Lebensverläufe und Wohlfahrtsentwicklung. Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1929–1931, 1939–*

- 1941, 1949–1951 (Teil I, S. 239–246). Berlin: Max-Planck-Institut für Bildungsforschung (Materialien aus der Bildungsforschung 35).
- Wagner, Michael. (1996a). *Ehestabilität und Sozialstruktur im Wandel der ost- und westdeutschen Gesellschaft*. Berlin: Freie Universität.
- Wagner, Michael. (1996b). Lebensverläufe und gesellschaftlicher Wandel. Die westdeutschen Teilstudien. *ZA-Information*, 38, 20–27.
- Wagner, Sandra J., & Visser, Karin. (2004). Die telefonische und schriftliche Nachrecherche zur Panelstudie „Ostdeutsche Lebensverläufe im Transformationsprozess (LV-Ost-Panel)“. Arbeitsbericht. In Anne Goedicke, Beate Lichtwardt, & Karl Ulrich Mayer (Hrsg.), *Dokumentationshandbuch Ostdeutsche Lebensverläufe im Transformationsprozess: LV-Ost-Panel* (S. 63–98). Berlin: Max-Planck-Institut für Bildungsforschung.
- Wetzler, S. E., & Sweeney, J. A. (1986). Childhood amnesia: An empirical demonstration. In David C. Rubin (Ed.), *Autobiographical memory* (pp. 191–201). Cambridge, UK: Cambridge University Press.
- White, R. T. (1982). Memory for personal events. *Human Learning*, 1, 171–183.
- Woolgar, Steve. (1983). Irony in the social study of science. In Karin Knorr-Cetina & Michael Mulkay (Eds.), *Science observed* (pp. 239–266). London: Sage.