

SISS:

Schriftenreihe

des Instituts für Sozialwissenschaften

der Universität Stuttgart: No. 3 / 2002

Probleme bei der Messung
individueller Veränderungsraten

13 empirisch und methodisch induzierte Effekte,
die es schwierig machen, Veränderungen von
generalisierten Bewertungen
zu ermitteln

Dieter Urban

**Institut für Sozialwissenschaften
Abteilung für Soziologie I
Universität Stuttgart**

70174 Stuttgart

Probleme bei der Messung individueller Veränderungsraten. 13 empirisch und methodisch induzierte Effekte, die es schwierig machen, Veränderungen von generalisierten Bewertungen zu ermitteln.

Z U S A M M E N F A S S U N G: Es werden verschiedene Faktoren genannt, die an der Entstehung von individuellen Veränderungen bei generalisierten Bewertungen und Urteilen beteiligt sein können und die methodisch kontrolliert werden sollten, da ansonsten die Messung von Veränderungsraten verzerrt werden kann. Dazu gehören sowohl soziale und kognitive Effekte, die die Veränderungen auf empirische Weise beeinflussen, die aber in der sozialwissenschaftlichen Analyse nicht ausreichend berücksichtigt werden, als auch methoden-induzierte Effekte, bei denen durch das Messverfahren selbst bestimmte Resultate der Messung entstehen. Insgesamt werden 13 solcher Effekte kurz erläutert und Hinweise zu ihrer empirischen oder statistischen Kontrolle gegeben.

Problems of measuring individual rates of change. 13 empirically and methodologically induced effects that make it difficult to detect changes of generalized value judgements.

A B S T R A C T: The research report presents various factors influencing individual changes of generalized value judgements that have to be controlled methodologically to avoid biased measurements of change rates. Among these factors are social and cognitive effects of the empirical processes that constitute individual judgements but that are often not being considered adequately in social science analysis. And there are methodologically induced effects that result from the process of measurement itself. The report briefly describes a total of 13 effects including some hints concerning their empirical and statistical control.

Im vorliegenden Bericht werden einige empirisch und/oder methodisch bedingte Faktoren benannt, die die Resultate von Längs- und Querschnittanalysen zur zeitlichen Veränderung von generalisierten subjektiven Bewertungen und Urteilsäußerungen, wie z.B. Veränderungen von Lebenszufriedenheiten, beeinflussen und auf diese Weise methodisch unerwünschte bzw. verzerrte oder verfälschte Messergebnisse erzeugen können.

Die hier aufgezeigten Probleme bei der Messung von Urteilsveränderungen entstehen entweder dadurch, dass die zur Analyse eingesetzten Mess- und Auswertungsinstrumentarien nicht in der Lage sind, empirisch relevante Regelmäßigkeiten, die bei der Entstehung von Urteilsveränderungen eine Rolle spielen, zu erkennen und damit spezifische Defizite zur Aufdeckung von Veränderungsprozessen aufweisen (a), oder dadurch, dass die eingesetzten Instrumentarien via ihrer Methodik selbständige Veränderungseffekte auslösen können (b), die dann nur noch schwerlich von den empirisch „wahren“ Veränderungsprozessen in der sozialen Lebenswelt zu unterscheiden sind.

Zu den empirisch induzierten Veränderungseffekten (a), die in der gängigen Forschung häufig übersehen werden, gehören:

- ! Kohorten-Effekte (vgl. E 2)
- ! dimensionale Effekte (vgl. E 3)
- ! situationsbedingte Extrem-Effekte (vgl. E 4)
- ! kognitiv-evaluativ bedingte Extrem-Effekte (vgl. E 5)
- ! evaluativ-bedingte Richtungs-Effekte (vgl. E 6)
- ! Homogenisierungs-Effekte (vgl. E 10)
- ! soziale Kontext-Effekte (vgl. E 12)

Zu den methodisch induzierten Veränderungseffekten (b), die messtechnisch und/oder auswertungsanalytisch verursacht werden können, gehören:

- ! Aggregations-Effekte (vgl. E 1)
- ! panel-bedingte Extrem-Effekte (vgl. E 7)
- ! panel-bedingte Negativ-Effekte (vgl. E 8)
- ! panel-bedingte Stabilitäts-Effekte (vgl. E 9)

! non-attitude-Effekte (vgl. E 10)

! Modell-Effekte (vgl. E 13)

Diese Liste erhebt natürlich keinen Anspruch auf Vollständigkeit. Es werden je nach Analyseperspektive auch noch weitere, hier nicht genannte Effekte, zu bestimmen sein, die die Messung von Veränderungseffekten in unerwünschter Weise beeinflussen können.

Im Folgenden werden Ursachen und Wirkungsweisen der oben benannten Effekte kurz erläutert sowie gegebenenfalls Möglichkeiten zu ihrer analytischen bzw. statistischen Kontrolle aufgezeigt.

E 1: Aggregations-Effekte

(Folge: zeit-ökologische Fehlschlüsse)

Die in der statistischen Auswertung zu beobachtende Variabilität von Bewertungsdaten wird oftmals bei einem Vergleich von mehreren, zeitlich verschobenen Aggregatanalysen (unabhängig davon, ob diese auf sequentiellen Querschnitts(Trend)-Erhebungen oder Panel-Erhebungen beruhen) eher gering ausfallen, obwohl die individuellen Variationen sehr hoch sein können. Als Folge können paradoxe Messergebnisse entstehen, bei denen hohe individuelle Instabilitäten und gleichzeitig hohe Aggregationsstabilitäten zu beobachten sind. Solche Ergebnisse werden z.B. aus der politischen Soziologie und dort vor allem aus der Wahl- und Partizipationsforschung berichtet (vgl. Kaase 1986).

Der Effekt ist deshalb so weit verbreitet, weil sich im Prinzip in der Aggregatanalyse stets gegenläufige individuelle Veränderungen neutralisieren können, so dass bei Anwendung aggregierender Auswertungsmethoden nur dann signifikante kollektive Veränderungsprozesse aufzuspüren sind, wenn diese einem typischen Veränderungsmuster entsprechen, das von einer deutlichen Mehrheit aller gemessenen, individuellen Variationen getragen wird.

Da deshalb in der Aggregatanalyse keine prinzipiell gültigen Aussagen über Richtung und Ausmaß von zeitabhängigen, individuellen Veränderungen erstellt werden können, und somit das Risiko „zeit-ökologischer Fehlschlüsse“ (Landua 1991) besonders groß ist¹⁾, müssen die

1) Ökologische Fehlschlüsse liegen immer dann vor, wenn von beobachteten Zusammenhängen auf Aggregatenebene in nicht erlaubter und somit verfälschender Weise auf das individuelle Verhalten einzelner Akteure geschlossen wird.

Vorteile der Panel-Analyse von Variabilitätsmessungen genutzt werden und individuelle Bewertungskarrieren untersucht werden, die jeweils nicht über alle Befragten aggregiert sind, sondern die getrennt für einzelne Ausgangslagen der Bewertung zu einem Referenzzeitpunkt t_0 aggregiert sind und nur dafür die jeweiligen Veränderungskarrieren thematisieren.

E 2: Kohorten-Effekte

(Folge: Muster des sozialen Wandels werden nicht bzw. falsch erkannt)

Auch wenn Aggregationseffekte (vgl. E1) aufgrund einer Panelanalyse vermieden werden können, ist es noch immer möglich, dass zeitliche Veränderungsprozesse falsch oder überhaupt nicht erkannt werden, weil sie kohortenspezifisch ablaufen: „Prozesse des sozialen Wandels vollziehen sich häufig über eine Differenzierung der Lebensläufe von Kohorten (Geburts-, Heirats-, Eintrittskohorten etc.), und können deswegen am besten über eine kohortenvergleichende Darstellung adäquat beschrieben werden.“²⁾

Das gilt in besonderer Weise für Veränderungen von Wertorientierungen, die häufig aufgrund von unterschiedlichen Sozialisationsbedingungen unterschiedlich stark intrapersonell verankert sind und damit auch unterschiedlich leicht variieren können (man denke z.B. an die Sozialisationshypothese im Inglehart'schen Wertwandel-Modell).

Ebenfalls ist zu berücksichtigen, dass individuelle Bewertungen an subjektiven Aspirationsniveaus ausgerichtet sind, welche in Abhängigkeit von der jeweiligen Lebensphase variieren können. So wurde z. B. gezeigt, dass mit zunehmendem Alter die allgemeine Lebenszufriedenheit deutlich ansteigt. Personen über 70 Jahre waren demnach zu fast einem Drittel mit ihrem Leben „hoch zufrieden“ (Landua 1993: 562).

Es sollten deshalb stets für mehrere Geburtskohorten zusätzliche Panelanalysen von Messwerten zu Bewertungsfragen durchgeführt werden. Erst eine Kombination von Längsschnitt- und Kohortenanalyse schafft die Voraussetzungen dafür, die Prozesse des sozialen Wandels zuverlässig erkennen zu können.

2) H.-P. Blossfeld in einer Buch-Rezension erschienen in: Soziologische Revue, 1994, No. 1, S. 189.

E 3: dimensionale Effekte (kognitive bzw. evaluative Auslöser vs. situative Auslöser)
(Folge: dimensional spezifische Veränderungsraten)

Die empirische Variabilität von Urteilsäußerungen ist generell hoch. Sie wird u.a. ausgelöst durch:

- a) Veränderungen von subjektiven Perzeptionsmustern (kognitiver Aspekt: Was wird wahrgenommen?),
- b) Veränderungen von subjektiven Bewertungsstandards (evaluativer Aspekt: Wie wird das Wahrgenommene bewertet?),
- c) Veränderungen von exogenen Lebensbedingungen (situativer Aspekt: Unter welchen individuellen Bedingungen wird was in welcher Weise bewertet?)

Ob für Veränderungen die Gründe vom Typ (a,b) oder vom Typ (c) relevant sind, kann allein dadurch festgestellt werden, dass die Analyse gruppenspezifisch durchgeführt wird und Gruppen unterschieden werden, für die sich bestimmte exogene Faktoren in spezifischer Weise verändert haben (z.B. als Veränderungen [nicht, mäßig oder stark] entsprechend ihrer Einkommenslage).

So konnte für den Einfluss von Arbeitslosigkeit (des Haushaltsvorstands) auf die Zufriedenheitswerte der einzelnen Familienmitglieder gezeigt werden, dass sich diese in besonders starker Weise auf Zufriedenheitsveränderungen bei den Kindern und am wenigsten auf Veränderungen bei den (Ehe-)Partnern auswirkt (vgl. Landua 1990b). Es muss also gruppentypische Variationen in den Bewertungsstandards externer Ergebnisse geben, die zu unterschiedlichen Evaluationen bei ansonsten konstanter materieller Betroffenheit führen.

In der Panel-Analyse sind deshalb auch Subgruppen zu bestimmen und auszuwerten, die sich nicht allein durch unterschiedliche Betroffenheitslagen, sondern durch unterschiedliche Framing-Prozesse unterscheiden. Anders wären differentielle, bidirektionale Wahrnehmungseffekte bei der Bewertung/Einschätzung exogener Faktoren nicht zu identifizieren.

ad: kognitive/evaluative Auslöser

Wenn die exogenen Lebensbedingungen sich wenig ändern, werden sich u.U. die subjektiven Bewertungen auch nur wenig oder überhaupt nicht ändern. Das ist durch die Analyse von zeit-homogenisierten Populationsgruppen zu erkennen, die hinsichtlich bestimmter, zu definierender sozio-demographischer Merkmale über alle Erhebungswellen hinweg konstant bleiben. Mögliche Homogenisierungsvariablen sind z.B.: Geschlecht (m/w), Arbeitslosigkeit (j/n), Krankhei-

ten (j/n), Ehe/Scheidung/Geburt (j/n).

In Fällen, wo relevante Lebenssituationen konstant bleiben und dennoch Veränderungen in Bewertungs- Urteilsmessungen auftreten, werden diese primär durch kognitive Neu-Bewertungen von an sich stabilen Lebensumständen ausgelöst sein. Neu-Bewertungen können eine kognitive Eigendynamik auch ohne sozio-externe Einflüsse in Gang setzen. Sie entstehen dann z.B.:

- (-) als Folge von Habitualisierung und Saturation. Dabei können z. B. Defizite bewusst werden, die ehemals nicht wahrgenommen wurden. Auch zuvor bereits bekannte Defizite können dabei in ihrem Ausmaß weitaus größer als zuvor erscheinen.
- (-) als Folge von Konsistenzanforderungen, die von anderen Wertorientierungen ausgehen und die in Richtung und Ausmaß von den zu messenden Urteilen abweichen. Dabei können zur Vermeidung von kognitiven Inkonsistenzen bzw. Dissonanzen und der damit verbundenen Stresserfahrungen unverändert gebliebene Lebenssituationen neu bewertet werden.³⁾

Mithin können geringe Veränderungen von subjektiven Bewertungen eher durch Wahrnehmungsveränderungen ausgelöst sein als durch Veränderungen exogener Lebensbedingungen.

Ein diesbezüglicher Test, der die Relevanz von Veränderungen kognitiver Orientierungen für die Verschiebung von subjektiven Bewertungsveränderungen misst, kann in den oben genannten, homogenisierten Subgruppen durchgeführt werden, indem dort Kausalmodelle spezifiziert und geschätzt werden, die als Prädiktoren neben den Ausprägungen von Kognitionen zu Beginn der Panel-Analyse (X_j) auch die Veränderungsraten dieser Kognitionen zwischen zwei verschiedenen Panel-Wellen (X_j) enthalten.

Z. B. analysieren Finkel/Schrott (1994) nicht nur den Einfluss von Links/Rechts-Selbsteinstufungen (LR) zum Zeitpunkt t_1 auf die Wahlentscheidung zum Zeitpunkt t_2 , sondern ihr diesbezügliches Regressionsmodell enthält als Prädiktor-Variable zusätzlich auch die Veränderung der Links/Rechts-Selbsteinstufung (LR) zwischen t_1 und t_2 .

ad: situative Auslöser

Dramatische Veränderungen der subjektiven Bewertungen werden in aller Regel nur durch

3) Die Bedeutung konsistenter kognitiver Orientierungen für die individuelle Lebenspraxis sozialer Akteure konnte z.B. die empirische Wahlforschung deutlich machen, welche solche Urteilsveränderungen, die in ihrer Richtung den bereits vor Durchführung von Wahlwerbe-Kampagnen bestehenden „prädispositionierten“ Parteien und/oder Kandidaten entsprachen, als besonders handlungsrelevant herausstellte (vgl. z.B. Finkel/Schrott 1994).

dramatische Veränderungen in den exogenen Lebensbedingungen ausgelöst. Zu den relevanten dramatischen Veränderungen gehören u.a. Veränderungen in den Bereichen von finanzieller Ausstattung, Beruf oder Familie.

Je direkter Bewertungsmessungen mit konkreten Lebensbedingungen verknüpft sind (im Unterschied zu generellen/universellen Einschätzungsmessungen), umso größer und zeitlich stringenter sind die ermittelten Variabilitäten. Gründe dafür dürften in der generell hohen Dynamik exogener gesellschaftlicher Faktoren liegen, die in der Wahrnehmung eher mit einer gewissen zeitlichen Verzögerung „eigenmächtig“ verzerrt werden können.

E 4: situationsbedingte Extrem-Effekte

(Folge: extreme Lebensbedingungen führen zu konstanten Bewertungen)

Die Relevanz situativer Auslöser (s.o.) unterliegt einem „situativen Extrem-Effekt“ (zu anderen Extrem-Effekten s. u.). Je extremer (positiv oder negativ) die exogenen Lebensbedingungen gekennzeichnet sind, umso konstanter (positiv oder negativ) erscheinen oft die subjektiven Bewertungen. Für diesen Effekt gibt es zwei mögliche Begründungen:

- a) Extreme Lebenssituationen weisen einen höheren Institutionalisierungsgrad als andere auf und sind deshalb dauerhafter, denn extreme Lebenssituationen unterliegen in stärkerem Ausmaß externen Einflüssen, die via individueller Entscheidung nicht zu modifizieren sind.
- b) Extreme Lebenssituationen besitzen eine größere „kognitive Prägekräft“, die somit auch zu einer höheren Permanenz der jeweils geprägten Bewertungen/Einschätzungen führt.

E 5: kognitiv-evaluativ bedingte Extrem-Effekte

(Folge: extreme Bewertungen werden weniger extrem)

Extreme Bewertungen (positiv oder negativ) unterliegen einem kognitiven Sättigungseffekt. Aufgrund ihrer Permanenz werden sie im Zeitverlauf die Bedeutung des Außergewöhnlichen verlieren und sich in Richtung weniger extremer Positionen verändern. Dabei wird es jedoch nicht zu sprunghaften Bewertungs/Einschätzungsveränderungen kommen (z. B. können ehemals unbedeutende, d. h. extrem irrelevante Defizite/Gratifikationen im Zeitablauf höhere Relevanzwerte erreichen).

E 6: evaluativ-bedingte Richtungs-Effekte

(Folge: negative Bewertungen sind weniger stabil)

Negative subjektive Bewertungen sind weniger stabil als positive Bewertungen. Von unzufriedenen Befragten erlebten z. B. 72 % eine positive Veränderung innerhalb eines Jahres, von zufriedenen Befragten nur 28 % (Landua 92: 226f). Für diesen Effekt sind zwei mögliche Gründe denkbar:

- a) Negative Bewertungen motivieren dazu, negative Lebensbedingungen zu verändern.
- b) Negative Bewertungen werden aufgrund der von ihnen provozierten kognitiven Dissonanzen eher umdefiniert als positive Effekte.

E 7: panel-bedingte Extrem-Effekte

(Folge: extreme Bewertungen werden weniger extrem)

Längsschnitt-Bewertungsmessungen unterliegen einem Skaleneffekt: extreme Positionen (negativ oder positiv) können sich entweder prinzipiell nicht mehr in Richtung des besetzten Extrems verändern, oder die Möglichkeiten einer weiteren Extremisierung sind weitaus geringer und damit auch unwahrscheinlicher als Veränderungen zum jeweiligen Gegenpol.

Messungen werden also künstlich stabiler, oder künstlich in Richtung der Skalenmitte verzerrt, wenn Befragte ein flexibles Antwortverhalten zeigen wollen (z. B. aufgrund von SD-Einflüssen).

Zur Überprüfung dieses Effekts müsste getestet werden:

- a) ob sich bei ansonsten vergleichbaren sozio-demographischen Randbedingungen die Messwerte von zunächst extrem bewertenden Befragten signifikant anders verändern als diejenigen solcher Befragten, die zunächst eher mittlere Bewertungen bevorzugten.
- b) ob bei Gültigkeit dieses Effektes in einer dritten Messwelle die Bandbreite der Antworten derjenigen Personen, die von „extrem“ nach „mittel“ gewandert sind, wieder zunehmen, da dann auch wieder extremere Positionen zu besetzen sind.

E 8: panel-bedingte Negativ-Effekte

(Folge: Bewertungen im Rahmen von schwach institutionalisierten Bewertungsmustern werden negativer)

Die mehrfach wiederholte Teilnahme an Befragungen zum gleichen Thema und mit den gleichen Messinstrumenten führt zu einem methodisch induzierten Antwortverhalten. Dies hat zur Konsequenz, dass in mäßiger, aber deutlicher Zunahme negative Bewertungswerte ermittelt werden.

Ursache für diesen Effekt können Gewöhnungs- und Lern-Einflüsse sein, die dann entstehen, wenn gleiche Bewertungen wiederholt abzugeben sind und damit eine Sensibilisierung für ein bestimmtes Thema erzeugt wird, die bei einer durchaus üblichen, latenten Alltags-Unzufriedenheit eher zu verstärkten Negativ-Bewertungen denn zur Zunahme von Positiv-Bewertungen führt.

Zur Kontrolle dieses Effekts wäre ein systematischer Vergleich von Erst- und Mehrfachbefragten durchzuführen.

E 9: panel-bedingte Stabilitäts-Effekte

(Folge: Bewertungen werden unabhängig von sozialen Dynamiken eher stabiler)

Stärker noch als panel-bedingte Negativ-Effekte (s.o.) wirken die methodisch ausgelösten Stabilitäts-Effekte (auch „Test-Retest-Effekte“ oder „socratic effects“ genannt). In Panel-Studien werden mit der Anzahl der Befragungen die gemessenen Werte signifikant und überdeutlich an Stabilität gewinnen, was sich z. B. am auffälligen Rückgang der Standardabweichungen bemerkbar machen kann (vgl. Landua 1993: 567) sowie in Kovarianz-Struktur-Modellen mit Mehr-Indikatoren-Messungen zur Erhöhung der Intra- und Interwellen-Konsistenz im Übergang von der ersten zur zweiten Erhebungswelle führt (vgl. Jagodzinski et al. 1987).

Wie bei allen Panel-Effekten werden auch hier Gewöhnungseinflüsse eine Rolle spielen, die zu einem sichereren Antwortverhalten führen können: bei wiederholten Befragungen werden die Befragten mit dem Messinstrument vertrauter und entdecken z. B. Ähnlichkeiten zwischen verschiedenen Items, was dann zu einer Erhöhung der Intra-Wellen-Konsistenz zwischen den Items eines bestimmten Messkonstrukts führen kann. Auch können im wiederholten Messprozess die Item-Inhalte für viele Befragte bedeutsamer und zentraler werden als sie es ur-

sprünglich waren bzw. überhaupt erst ihre Bedeutung und Zentralität erhalten, da mit der Befragung ein kognitiver Verarbeitungsprozess ausgelöst wird (von dem wiederum auch zusätzliche Konsistenzanforderungen für die Beantwortung inhaltlich verwandter Fragen ausgehen werden).

Zudem wirken die hier separat beschriebenen panel-bedingten Extrem-Effekte tendenziell stabilitäts-verstärkend: wer zunächst extrem geantwortet hat und dann zusätzliche Verbesserungen/Verschlechterungen erfährt, kann nur bei seiner Extrem-Bewertung verbleiben bzw. diese nur relativ geringfügig radikalieren.

Zur Kontrolle der hier beschriebenen Effekte müsste ein systematischer Vergleich von Erst- und Mehrfachbefragten durchgeführt werden. Zusätzlich kann eine Mehr-Indikatoren/Mehr-Wellen/Kovarianz-Strukturanalyse darüber Aufschluss geben, inwiefern im Zeitverlauf zwischen den Konstrukten und innerhalb bzw. zwischen den dazu gemessenen Indikatoren bedeutungsvolle Konsistenzveränderungen auftreten. In systematisierter Form betrachtet, können sie im Rahmen von vier Konsistenz-Typen analysiert werden:

- a) als Konsistenzsteigerung der Beziehungen zwischen den latenten Konstrukten über die Zeit hinweg („growing intertemporal latent consistency“, Jagodzinski et al. 87:262)
- b) als Konsistenzsteigerung der wellen-spezifischen Beziehungen zwischen Konstrukt und Messwert über die Zeit hinweg („growing intrawave latent consistency“, a.a.O.)
- c) als Konsistenzsteigerung der Beziehungen zwischen den zeitspezifischen Messungen aller Indikatoren für ein Konstrukt über die Zeit hinweg („increasing intrawave observed consistency“, a.a.O.)
- d) als Konsistenzsteigerung der Beziehungen zwischen den Messungen eines bestimmten Items über die Zeit hinweg („increasing interwave observed consistency“, a.a.O.).

E 10: Homogenisierungseffekte

(Folge: unterschiedliche Veränderungsraten innerhalb einer Gruppe sind nicht feststellbar)

Bei der schriftlichen Befragung von Gruppenmitgliedern (z.B. den Mitgliedern einer Familie) kann es aufgrund von Kontroll- und SD-Einflüssen (z.B. bei gemeinsamer Beantwortung des Fragebogens) zur Anpassung/Homogenisierung von Messwerten kommen, obwohl die empirischen Werte von Ausprägungsintensitäten und Veränderungsausmaßen der gemessenen

Bewertungen und Urteils zwischen den Mitgliedern einer Familie durchaus unterschiedlich sind.

So konnten z. B. die statistisch signifikanten Differenzen zwischen Zufriedenheitsmessungen bei zwei verschiedenen Stichproben (Wohlfahrtssurvey u. Sozio-ökonomisches Panel) zum Teil auf unterschiedliche Untersuchungseinheiten zurückgeführt werden (Einzel-Befragung vs. Haushaltsbefragung): bei Beschränkung der Analyse auf die Alleinwohnenden verschwand die zuvor beobachtete Differenz in den Messwert-Verteilungen (vgl. Mohr 1987).

Auch werden Homogenisierungs-Effekte nach den Ergebnissen der methodologischen Forschung vor allem dann auftreten, wenn die „wahren“ Werte nur schwach ausgeprägt sind oder eher von „non-attitudes“ auszugehen ist (vgl. non-attitude-Effekte).

In diesem Sinne kann eine Mehrwellen-Panelstudie durchaus mit einer politischen Wahlkampf-Kampagne verglichen werden, für die sich auch herausgestellt hat, dass die öffentliche Bewusstmachung von politischen Themen zu einer Verstärkung des Konformitätsdrucks subjektiv relevanter, sozialer Netzwerke führt: Derjenige, für den ein extern vorgegebenes Thema neu ist, wird zunächst im unmittelbaren sozialen Umfeld nach Orientierung suchen.

Anhaltspunkte zur empirischen Relevanz des Homogenisierungs-Effektes können Vergleiche zwischen den Messwerten unterschiedlich komplexer Untersuchungseinheiten, wie z.B. denjenigen von Alleinlebenden und Familienmitgliedern, bei ansonsten gleichen sozio-demographischen Ausprägungen erbringen. Allerdings stehen dazu oftmals nicht genügend Fallzahlen für jeden Untersuchungstyp zur Verfügung (z.B. gibt es in der Regel zu wenige Informationen von Alleinlebenden).

Ohne gruppenspezifische Vergleichsinformationen (z.B. von Alleinlebenden) lässt sich ein Test nutzen, bei dem zwei geschachtelte Kausalmodelle geschätzt und verglichen werden. Dabei enthält das reduzierte Modell allein die Prädiktoren-Messwerte der jeweiligen Analysegruppe (z. B. von Vätern, Müttern oder Kindern), während das umfassendere Kausalmodell zusätzlich die gleichen Prädiktoren-Messwerte auch von der jeweiligen Gruppe enthält, von der angenommen wird, dass von dieser ein Homogenisierungseffekt ausgeht.

In einer australischen Studie (Hayes/Bean 1992) wurden mit diesem Test fünf verschiedene Modelle untersucht, die jeweils eine bestimmte Werthaltung über den Einfluss von 8 Prädiktoren erklären sollten. Wurde jedes der fünf Modelle um die jeweiligen 8 Prädiktoren der Partnergruppe erweitert (d. h. also: das „Männer-Modell“ um die gleichen Prädiktoren der zugehörigen weiblichen Partner und umgekehrt), veränderte sich in beiden Geschlechtskonstellationen bei

4 von 5 Modellen der Anteil erklärter Varianz (adjusted) in signifikanter Weise. Dabei lag R^2 im erweiterten Modell durchschnittlich um ca. 10 % höher als im reduzierten Modell. Allerdings veränderten sich durch die Erweiterung der Modelle die Koeffizientenschätzwerte nicht derart dramatisch, dass dies zu neuen substantiellen Interpretationen der Resultate geführt hätte. Jedoch sind solche Veränderungen modell- und stichprobenspezifisch und können im jeweiligen Anwendungsfall durchaus inhaltlich bedeutsame Ausmaße annehmen.⁴⁾

Weitere Tests hinsichtlich der Bedeutung von Homogenisierungseffekten können durch Anpassung der für die Überprüfung von sozialen Kontext-Effekten beschriebenen Methoden durchgeführt werden.

E 11: non-attitude-Effekte

(Folge: Bewertung/Einstellungsmessungen werden hochgradig invalide)

In der politischen Einstellungsforschung haben sich viele Messergebnisse als die Beobachtung von so genannten Nicht-Einstellungen oder „non-attitudes“ erwiesen. Messungen von non-attitudes liegen immer dann vor, wenn auf Fragestimuli allein zufallsgesteuert geantwortet wird, z.B. aufgrund von „ad-hoc-feelings“, die während der Befragung entstehen, die aber dann dennoch dazu führen, dass dezidierte Einschätzungen zu Bewertungsfragen (etwa: „stimme voll oder ganz zu“) abgegeben werden. Non-Attitude-Messungen können aber auch dadurch entstehen, dass Antworten durch Einflüsse einer Fremddimension zustande kommen (z. B. aufgrund von SD-Einflüssen).

Converse (1970, 1974) erzielte bei diversen Einstellungsmessungen mit je drei Erhebungszeitpunkten nur Test-Retest-Korrelationen zwischen 0.23 und 0.46 sowie Korrelationen zwischen den ersten und dritten Messzeitpunkten, die nicht (wie eigentlich zu erwarten) wesentlich niedriger als diejenigen zwischen zwei benachbarten Messzeitpunkten waren. Converse geht deshalb in seinem diesbezüglichen Erklärungsmodell davon aus, dass es stets zwei Gruppen von Einstellungsträgern gibt: diejenigen, deren Werte über die Zeit hinweg stabil bleiben (und die feste Übergangsraten in einer $t_1 - t_2$ Matrix aufweisen), und diejenigen, deren Werte sich rein zufällig (stochastisch) verändern (mover-stayer-Modell).

4) Die Ergebnisse des beschriebenen Vergleichstests sollten zusätzlich auf jeden Fall hinsichtlich zu erwartender Multikollinearitätseffekte überprüft werden.

Bei Frage-Items mit hohem non-attitude-Potential können nach den Ergebnissen von Converse etwa 20 % bis 35 % aller Befragten als solche „mover“ klassifiziert werden, deren Einstellungswerte entweder nur ein sehr geringes Ausmaß von kognitiver Zentralität aufweisen, oder die keine hinreichenden Informationen besitzen, um stabile und substanzielle Einstellungswerte wiedergeben zu können. Damit wären natürlich auch darauf basierende Veränderungsmaße zum sozialen Wandel von Einstellungen ohne Gültigkeit.

Um non-attitude-Effekte in der Einstellungsanalyse zu verhindern, sollte die Zentralität von einstellungsrelevanten Überzeugungen (beliefs) getestet werden. Dazu müssten zunächst diejenigen Items identifiziert werden, die potenziell „non-attitude-verdächtig“ sind. Als dementsprechender Indikator kann die Rate der „weiß nicht“ oder „k. A.“ Antworten benutzt werden (die bei den Converse-Daten zwischen 29 und 36 % lag).

Eine weitere Möglichkeit, „non-attitude-verdächtige“ Fragen-Items zu identifizieren, wäre dann gegeben, wenn Fragen zum gleichen Zeitpunkt leicht modifiziert (z. B. durch Drehung der Skalen-Pole) wiederholt wurden, und die Antworten nur sehr schwache Korrelationen aufweisen.

Bei „non-attitude-verdächtigen“ Frage-Items sollten die Wechsler zwischen t_1 und t_2 dadurch getestet werden, dass ihre Antworten zu t_3 als Kriterium für einen wahren Wechsel zu t_2 benutzt werden, d.h. es sollte die Korrelation zwischen t_2 und t_3 darüber entscheiden, ob es sich um Zufallsveränderungen handelt oder nicht (dabei muss natürlich als Annahme unterstellt bzw. getestet werden, dass es im sozialen Feld keinen Grund für einen Einstellungswandel gab).

Im Falle eines begründeten Verdachts auf „non-attitude-verzerrte“ Messdaten, kann die Analyse dadurch entzerrt werden, dass nur diejenigen Befragten in der Untersuchung berücksichtigt werden, bei denen feste Antwortmuster zu thematisch verwandten, anderen Einstellungsthemen vorliegen. Damit könnte allerdings nicht ausgeschlossen werden, dass sich diese Stabilität allein aufgrund des übermächtigen Einflusses einer Fremddimension (z. B. aufgrund von familiären SD-Effekten) ergibt. Denn Personen ohne „wahren“ Messwert sind besonders leicht von zweiter oder dritter Seite zu beeinflussen. Zudem wären alle in diesem Sub-Sample erzielten Ergebnisse „choice-based“ und somit inferenzstatistisch nur unter zusätzlichen Annahmen zu testen.

E 12: soziale Kontext-Effekte

(Folge: 1. invalide Messungen individueller Veränderungen
2. invalide Schätzung individueller Veränderungseffekte)

Die Veränderung von individuellen Bewertungswerten kann von sozialen Kontextbedingungen abhängig sein. Z.B. ist dann von familiären Kontexteffekten auszugehen, wenn die interessierenden Urteilsveränderungen aller Familienmitglieder in gleicher Richtung und in gleichem Umfang erfolgen. Dementsprechend sind familiäre Kontexteffekte dann unwahrscheinlich und individuelle Effekte wahrscheinlich, wenn die betreffenden Wert- und Urteilshaltungen von geringer familiärer Institutionalisierung sind. In diesen Fällen werden zentrale Kognitionen/Evaluationen nicht von allen Mitgliedern der Familie geteilt und sind möglicherweise Ausdruck von sich überschneidenden, sozialen Netzwerken mit unterschiedlichen kognitiven Ausrichtungen.

Zusätzlich können daraus entspringende Konflikte zwischen rivalisierenden substanziellen Orientierungen und sozialen Konformitätsverpflichtungen zu einer erhöhten Dynamik von Urteilsveränderungen führen.

Das Verhältnis und die relative Bedeutung von individuellen versus kontextuellen Effekten bzw. Veränderungen kann im Rahmen von Strukturmodellen getestet werden, die sowohl Komponenten enthalten, deren Varianzen zwischen den Kontexten variieren (z. B. zwischen Familien), als auch solche Komponenten, deren Varianzen innerhalb von Kontexten variieren (die z. B. spezifisch für verschiedene Familienmitglieder sind). Im Rahmen solcher Modelle lässt sich entscheiden:

- a) ob soziale Aggregate, wie z. B. Familien, typische Veränderungsmuster aufweisen und deshalb als homogene Einheiten analysiert werden sollten,
- b) ob und wenn ja, wie stark und wie unterschiedlich Kontextmerkmale die individuellen Elemente des Kontextes (z. B. Familienmitglieder) beeinflussen,
- c) ob die evtl. vorhandene Relevanz von Kontexteffekten einen Einfluss auf die Stärke evtl. vorhandener Individualeffekte hat (was z. B. das Verhältnis von interfamiliären versus intrafamiliären Differenzen bei Urteilsveränderungen und deren Verursachungen betrifft).

Wenn zur Konstruktion der oben beschriebenen Modelle keine direkten kontextbezogenen Messwerte vorliegen, so können diese in Form eines zweistufigen Kovarianz-Struktur-Modells simuliert werden. Latente Kontextfaktoren sollten sich darin auf alle Faktoren beziehen, die

spezifisch für die speziellen Kontext-Komponenten angesetzt wurden (z. B. Väter, Mütter und Kinder), und beeinflussen somit im Messmodell alle Indikatoren, die dem jeweiligen Kontext zuzurechnen sind (vgl. dazu Hauser/Mossel 1987).⁵⁾

Weitere Tests hinsichtlich der Bedeutung von sozialen Kontext-Effekten können durch Anpassung der für die Überprüfung von Homogenisierungseffekten beschriebenen Methoden durchgeführt werden.

E 13: Modell-Effekte

(Folge: Fehldiagnose, d. h. Analyse von nicht-vorhandenen bzw. Nicht-Analyse von vorhandenen substantziellen Veränderungen)

Signifikante Veränderungen von empirischen Indikator-Werten über eine oder mehrere Erhebungswellen hinweg können nicht ohne weiteres als Indiz für substantziell relevante Veränderungen von theoretisch interessanten Sozialformen benutzt werden. Zeitliche Veränderungen in Indikator-Variablen können auch eine Folge veränderter Mess-Reliabilitäten oder veränderter Fehler-Varianzen sein.

Ob eine empirisch beobachtete Veränderung tatsächlich auf Urteilsveränderungen zwischen den Erhebungswellen schließen lässt, kann nur durch die Spezifikation und Schätzung von Kovarianz-Strukturmodellen mit Mehr-Indikatoren-Messungen ermittelt werden (vgl. Jagodzinski et al. 1987; Raffalovich/Bohrnstedt 1987). In diesen Modellen ist jeder Messwert bestimmt:

- a) vom Einfluss eines für mehrere verwandte Mess-Items gemeinsamen Faktors („common factor“), der in der Regel dem theoretisch interessierenden Konstrukt entspricht (und sich in der Zeit verändert oder auch nicht),
- b) vom Einfluss eines speziellen Faktors („unique factor“), der für jedes Mess-Item spezifisch ist und deshalb zu Autokorrelationen zwischen den Messwerten für ein Item über alle Wellen hinweg führen kann,
- c) vom Einfluss einer Fehlergröße, die entweder als klassischer Zufallsfehler oder als spezifischer Fehler (z. B. als Panel-Responsefehler für best. Panel-Wellen) modelliert werden kann und auch zu Autokorrelationen führt (z.B. bei Zufallsfehlern aufgrund von Erinnerungs-

5) Leider weisen derartige Modelle erhebliche Identifikationsprobleme auf, die in der Regel nur durch ein Bündel „schmerzhafter“ Restriktionen zu beseitigen sind.

effekten).

Literatur

- Converse, P.E., 1970: Attitudes and non-attitudes: continuation of a dialogue. S. 168-189 in: Tuftte, E. (ed.), *The quantitative analysis of social problems*. Reading: Addison-Wesley.
- Converse, P.E., 1974: "Comment: The status of non-attitudes. *American Political Science Review* 68: 650-660.
- Finkel, S.E./Schrott, P., 1994: Wählerstimmen durch Wahlkämpfe? Eine Analyse der Bundestagswahl 1990. *ZUMA-Nachrichten* 18: 7-34.
- Hauser, R.M./Mossel, P.A., 1987: Some structural equation models of sibling resemblance in educational attainment and occupational status. S. 108-137 in: Cuttance, P./ Ecob, R. (eds.), *Structural Modeling by Example*. New York: Cambridge University Press.
- Hayes, B./Bean, C.S., 1992: The Impact of Spousal Characteristics on Political Attitudes in Australia. *Public Opinion Quarterly* 56: 524-529.
- Jagodzinski, W. et al., 1987: Is There a "Socratic Effect" in Nonexperimental Panel Studies? Consistency of an Attitude Toward Guestworkers. *Sociological Methods and Research* 15: 259-302.
- Kaase, M., 1986: Das Mikro-Makro-Puzzle der empirischen Sozialforschung. Anmerkungen zum Problem der Aggregatsstabilität bei individueller Instabilität in Panelbefragungen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 38: 209-222.
- Landua, D., 1990: Erweiterungsmöglichkeiten der Standardverfahren der empirischen Sozialforschung. *Panelauswertungen von Parteipräferenzen*. Berlin: WZB.
- Landua, D., 1991: Möglichkeiten längsschnittorientierter und kontextbezogener Auswertungsverfahren für die Analyse subjektiver Einstellungsdaten. *Politische Vierteljahresschrift* 32: 92-110.
- Landua, D., 1993: Veränderungen von Zufriedenheitsangaben in Panelbefragungen. Eine Analyse über nicht beabsichtigte Effekte des Längsschnittdesigns. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 45: 553-571.
- Mohr, H.M., 1987: Analysen zur Vergleichbarkeit von Zufriedenheitsmessungen. *Zeitschrift für Sozialpsychologie* 18: 160-168.
- Raffalovich, L.E./Bohrstedt, G.W., 1987: Common, Specific and Error Variance Components of Factor Models: Estimation with Longitudinal Data. *Sociological Methods & Research* 15: 385-405.

**SISS: Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart
bisher sind erschienen:**

- No. 1/1994 "Vertrauen" - soziologisch betrachtet. Ein Beitrag zur Analyse binärer Interaktionssysteme.
Peter Antfang, Dieter Urban
- No. 2/1994 Report on the German Machine Tool Industry.
Frank C. Englmann, Christian Heyd, Daniel Köstler, Peter Paustian
with the assistance of Susanne Baur and Peter Bergmann
- No. 3/1994 Neue württembergische Rechtstatsachen zum Unternehmens- und Gesellschaftsrecht.
Udo Kornblum
- No. 4/1994 Rechtstatsachen zum Unternehmens- und Gesellschaftsrecht aus den neuen Bundesländern.
Udo Kornblum
- No. 1/1995 Die Bedeutung Neuronaler Netze in der Ökonomie.
Hermann Schnabl
- No. 2/1995 Regionale Strukturprobleme.
Sammelband der Beiträge zum Symposium vom 13. und 14. Oktober 1994.
Frank C. Englmann (Hrsg.)
- No. 3/1995 Latent Attitude Structures Directing the Perception of New Technologies.
An Application of SEM-Methodology to the Construction of Attitude Measurement Models Related to Technologies of Prenatal Genetic Engineering and Testing.
Dieter Urban
- No. 4/1995 Handbuch zur empirischen Erhebung von Einstellungen/Kognitionen zur Bio- und Gentechnologie (inklusive Diskette; zweite, überarbeitete und erweiterte Auflage).
Uwe Pfenning, Dieter Urban, Volker Weiss
- No. 5/1995 Social Indicators in a Nonmetropolitan County: Testing the Representativeness of a Regional Nonrandom Survey in Eastern Germany.
Dieter Urban, Joachim Singelmann
- No. 1/1996 Jugend und Politik im Transformationsprozeß. Eine Fallstudie zur Stabilität und Veränderung von politischen Einstellungen bei ostdeutschen Jugendlichen zwischen 1992 und 1995.
Dieter Urban, Joachim Singelmann, Helmut Schröder
- No. 2/1996 Einstellungsmessung oder Einstellungsgenerierung? Die Bedeutung der informationellen Basis bei Befragten für die empirische Rekonstruktion von Einstellungen zu gentechnischen Anwendungen.
Martin Slaby
- No. 1/1997 Gentechnik: „Fluch oder Segen“ versus „Fluch und Segen“. Bilanzierende und differenzierende Bewertungen der Gentechnik in der öffentlichen Meinung.
Dieter Urban und Uwe Pfenning

(Fortsetzung)

**SISS: Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart
bisher sind erschienen (Fortsetzung):**

- No.2/1997 Die soziale Vererbung von Ausländer“feindlichkeit“. Eine empirische Längsschnittanalyse der intra- und intergenerativen Transmission von sozialen Einstellungen.
Dieter Urban und Joachim Singelmann
- No. 3/1997 Politische Sozialisation im Transformationsprozeß: Die Entwicklung demokratiebezogener Einstellungen von ostdeutschen Jugendlichen und deren Eltern zwischen 1992 und 1996.
Barbara Schmidt, Dieter Urban, Joachim Singelmann
- No.1/1998 Bewertende Einstellungen zur Gentechnik: ihre Form, ihre Inhalte und ihre Dynamik. Kurzbericht zu Ergebnissen des Forschungsprojektes „Einstellungen zur Gentechnik“.
Dieter Urban, Uwe Pfenning, Joachim Allhoff
- No.2/1998 Technikeinstellungen: gibt es die überhaupt?
Ergebnisse einer Längsschnittanalyse von Bewertungen der Gentechnik.
Dieter Urban
- No.3/1998 Zur Interaktion zwischen Befragten und Erhebungsinstrument.
Eine Untersuchung zur Konstanz des Meinungsurteils von Befragten im Interviewverlauf.
Martin Slaby
- No.1/1999 Role Models and Trust in Socio-Political Institutions:
A Case Study in Eastern Germany, 1992-96.
Joachim Singelmann, Toby A. Ten Ayck, Dieter Urban
- No.1/2000 Die Zufriedenheit von Stuttgarter Studierenden mit ihrer Lebens- und Wohnsituation. Erste deskriptive Ergebnisse einer sozialwissenschaftlichen Studie zu allgemeinen und bereichsspezifischen Zufriedenheiten der Studierenden des Campus Vaihingen und des Campus Hohenheim.
Projektgruppe Campus: Slaby, M., Grund, R., Mayerl, J., Noak, T., Payk, B., Sellke, P., Urban, D., Zudrell, I.
- No.2/2000 Längsschnittanalysen mit latenten Wachstumskurvenmodellen in der politischen Sozialisationsforschung.
Dieter Urban
- No.1/2001 Unser „wir“ - ein systemtheoretisches Modell von Gruppenidentitäten.
Jan A. Fuhse
- No.2/2001 Differentielle Technikakzeptanz, oder: Nicht immer führt die Ablehnung einer Technik auch zur Ablehnung ihrer Anwendungen.
Eine nutzentheoretische und modell-statistische Analyse.
Martin Slaby, Dieter Urban
- No.3/2001 Religiosität und Profession. Longitudinale Analysen zur Entwicklung des religiösen Selbstbildes bei Erzieherinnen.
Heiko Lindhorst
- No.4/2001 Ist Glück ein affektiver Sozialindikator subjektiven Wohlbefindens?
Dimensionen des subjektiven Wohlbefindens und die Differenz zwischen Glück und Zufriedenheit.
Jochen Mayerl

(Fortsetzung)

**SISS: Schriftenreihe des Instituts für Sozialwissenschaften der Universität Stuttgart
bisher sind erschienen (Fortsetzung):**

- No.1/2002 Risikoakzeptanz als individuelle Entscheidung.
Zur Integration der Risikoanalyse in die nutzentheoretische
Entscheidungs- und Einstellungsforschung.
Martin Slaby, Dieter Urban
- No.2/2002 Vertrauen und Risikoakzeptanz. Zur Relevanz von Vertrauen
bei der Bewertung neuer Technologien.
Martin Slaby, Dieter Urban
- No.3/2002 Probleme bei der Messung individueller Veränderungsraten.
13 empirisch und methodisch induzierte Effekte, die es schwierig machen,
Veränderungen von generalisierten Bewertungen zu ermitteln.
Dieter Urban