

Wie reliabel sind Fragen zur Kriminalität und Kriminalitätsfurcht?

Ergebnisse einer Test-Retest-Studie

von Karl-Heinz Reuband und Parviz Rastampour

Zusammenfassung

Die Reliabilität der Messung der Einstellungen zur Kriminalität, der Furcht der Kriminalität und der Viktimisierung ist in deutschen Untersuchungen wenig erforscht. Die vorgestellte Studie untersucht die Höhe der Reliabilität entsprechender Indikatoren im Vergleich mit Indikatoren zur Erfassung anderer Aspekte der alltäglichen Realität (z.B. Fragen der Wirtschaft), das Verhältnis der Reliabilitäten auf Individual- und Aggregatebene zueinander und die Bestimmungsfaktoren der Reliabilität auf der Basis einer durch technische Fehler zufälligerweise doppelt postalisch befragten Teilpopulation (N = 45), die sich von der Gesamtstichprobe in den relevanten Parametern nicht signifikant unterscheidet. Die Reliabilitäten sind zufriedenstellend und unterscheiden sich für den Bereich Kriminalität (Mittelwert .67) nicht wesentlich von denen für den Bereich Wirtschaft (.60). Insgesamt zeigt sich, daß für den Bereich Kriminalität die Motivation zur Mitarbeit für die Reliabilität entscheidend ist. Diese gilt es durch gezielte Anlage des Erhebungsinstrumentes und seine Anwendung zu optimieren.

Abstract

As research about reliability of instruments for the measurement of attitudes towards crime, fear of crime and victimization is highly needed in Germany, the authors explore the reliability of indicators of those constructs in comparison with indicators measuring attitudes to economic aspects, the relationship between reliabilities on aggregate and individual level, and predictors of reliability based on a small sample (N= 45) of respondents who answered the mail questionnaire two times because of a technical error. There were no significant differences between this small sample and the original study population with respect to central socio-demographic parameters. Reliabilities found are satisfactory and there are no relevant differences between crime and economy items. Motivation to participate in the survey is a central precondition to measure crime indicators reliably and must be optimized by the design of the questionnaire and its application.

1. Einleitung

Die Zahl der Bevölkerungsumfragen zur Viktimisierung und zur Kriminalitätsfurcht wächst und mit ihnen auch die Erkenntnis, wie sehr methodische Aspekte bei der Interpretation berücksichtigt werden müssen. Denn ein gleichartiges Vorgehen sichert keineswegs zwangsläufig gleichartige empirische Befunde. Indikatoren, die sich auf identische Sachverhalte beziehen und lediglich etwas anders formuliert sind – sei es auf der Ebene der Fragen oder der Antwortkategorien –, erbringen oftmals recht unterschiedliche Antwortverteilungen (vgl. u.a. Reuband 2000a). Und selbst dort, wo identische Indikatoren in mehreren, nahezu zeitgleich durchgeführten Erhebungen eingesetzt werden, kommen z.T. erhebliche Unterschiede vor (Schnell/Kreuter 2000). Einige dieser Diskrepanzen mögen Folge des jeweiligen Fragekontextes sein, in dem die Fragen gestellt wurden. In anderen Fällen dürften es Unterschiede in der Praxis der Stichprobenziehung sein, und in wieder anderen Fällen der Modus der Datenerhebung. So erbringen z.B. telefonische Erhebungen niedrigere Werte für Kriminalitätsfurcht und Viktimisierung als postalische Befragungen (vgl. z.B. Reuband 2000b).

Welchen Anteil die Indikatoren als solche an den beschriebenen Effekten haben, – wie sehr sie selbst auch Ausdruck instabiler Einstellungen im Interview sein können – ist bislang ungeklärt. Nach Untersuchungen von Philip Converse (1974, 1970) sind viele Einstellungen der Bürger – vor allem zu komplexeren politischen Themen – wenig reflektiert und wenig auskristallisiert. Ihre Antworten sind entsprechend wenig reliabel und schwanken über die Zeit. Sie antworten mehr oder minder zufällig. Beim Thema der Kriminalität ist eine solche Tendenz nicht zwangsläufig zu erwarten, da es normalerweise in erheblichem Maße das Alltagsbewußtsein des Bürgers prägt und für den einzelnen subjektiv zentral ist. Subjektive Zentralität konstituiert nach Converse gewöhnlich eine bedeutsame Voraussetzung für Stabilität auf der issuebezogenen Ebene. Unabhängig vom Aspekt der themenspezifischen Einstellungsstruktur fragt sich, inwieweit die von Converse konstatierte Instabilität der Antworten auch methodisch bedingt sein könnte – als eine Folge der jeweiligen Befragungssituation: in den von ihm untersuchten Fällen handelt es sich um face-to-face-Befragungen. Diese aber unterliegen im besonderen Maße der Gefahr situationsspezifischer Antworttendenzen, die u.a. durch Anwesenheit des Interviewers und die Anwesenheit dritter Personen hervorgerufen werden können.

2. Zielsetzung und Methoden

Im folgenden soll über eine Untersuchung berichtet werden, in der die Reliabilität im Rahmen wiederholter Messung für Fragen der Kriminalitätserfahrung, Kriminalitätsfurcht und der Einstellungen zur Kriminalität erhoben wurde. Uns interessieren drei Fragen: Wie groß ist die Reliabilität der Indikatoren im Vergleich zu Indikatoren, die sich auf andere Themen des Alltags beziehen? Wie sehr verschieben

sich die Einstellungen bei wiederholter Messung auf der Aggregat- und auf der Individualebene? Und welche Einflußfaktoren bestimmen über die Höhe des Reliabilitätskoeffizienten? Die Untersuchung stützt sich auf eine postalische Erhebung. Postalische Befragungen haben den Vorteil, daß sozial erwünschte Antworten seltener gegeben werden. Und sie haben den Vorteil, daß der Befragte mehr Zeit hat, die Antwort zu überlegen. Zufällige, mehr oder minder durch die Situation geprägte ad hoc-Antworten dürften seltener vorkommen.

Die Untersuchung basiert auf einer repräsentativen Randomstichprobe der wahlberechtigten Bevölkerung der Städte Dresden, Chemnitz und Leipzig (18 Jahre und älter). Die Befragung wurde in Anlehnung an die Total Design Method von Dillman mit bis zu drei Mahnaktionen durchgeführt (Dillman 1978).¹ Für 45 der rund 3600 Befragten liegen doppelte Fragebögen vor. Sie liegen doppelt vor, weil die Befragten den gleichen Fragebogen mehrfach erhielten und auch mehrfach ausfüllten. Dies geschah entweder, weil sie nach dem Ausfüllen des Fragebogens eine Mahnung mit einem weiteren Fragebogen erhalten hatten, während der erste ausgefüllte Fragebogen noch unterwegs war. Oder es geschah lange nachdem der Fragebogen ausgefüllt zurückgesandt worden war und aufgrund eines technischen Irrtums Befragte erneut angeschrieben wurden – in der Annahme, sie hätten noch nicht geantwortet. Nach den verfügbaren Rückmeldungen glaubten viele Teilnehmer daraufhin, ihr Fragebogen sei verloren gegangen und sie müßten erneut antworten. Das Vorhandensein doppelt ausgefüllter Fragebögen ermöglicht uns eine Überprüfung der Reliabilität der Antworten. Wir können anhand der doppelten Fragebögen, die eindeutig von denselben Personen ausgefüllt worden sind, überprüfen, inwieweit die Antworten beim zweiten Mal von denen der ersten abweichen.

Die Befragten, für die Fragebögen zweifach vorliegen, unterscheiden sich nicht nennenswert in der sozialen Zusammensetzung von der Gesamtheit der Befragten. Mit einer Ausnahme: Personen aus Chemnitz sind überrepräsentiert, weil dort aus technischen Gründen mehr Zielpersonen als in den beiden anderen Städten mehrfach angeschrieben wurden, obwohl sie bereits geantwortet hatten. Natürlich handelt es sich bei dieser Studie nicht im stringenten Sinne um eine Test-Retest Studie, bei der alle Befragten in eine zweite Messung einbezogen werden. Selektionseffekte bei den Befragten sind möglich, aber in ihren Konsequenzen vermutlich wenig bedeutend. Es ist nicht anzunehmen, daß es sich bei den Befragten um Personen handelt, denen es besonders um die Vermittlung ihrer eigenen Einstellungen ging und die deshalb besonders hohe oder geringe Reliabilitäten aufweisen. Wahrscheinlicher ist, daß wir es de facto mit einer Art Zufallsstichprobe der Befragten zu tun haben. Gleichwohl: Die Studie ist ein erster Schritt in Richtung auf eine Test-Retest-Studie hin, die für die Kriminalitätsthematik bislang fehlte. Die Ergebnisse sind daher als tentativ zu beurteilen.

Test-Retest-Studien sollten zwischen den Erhebungswellen im allgemeinen einen kurzen Zeitabstand aufweisen, um periodenbedingte Effekte zu reduzieren. Sie sollten andererseits auch nicht in zu kurzem Abstand erfolgen, um zu verhindern,

daß sich die Befragten an ihre zuletzt gegebenen Antworten erinnern (vgl. auch Bortz und Döring 1995). Im vorliegenden Fall liegt der zeitliche Abstand zwischen dem Ausfüllen der Fragebögen bei durchschnittlich 50 Tagen. Dies ist ein relativ langer Zeitraum für Test-Retest-Studien. Eine Rückerinnerung an die einst gegebenen Antworten ist jedoch nicht nur aufgrund dieses relativ langen Zeitraums unwahrscheinlich. Sie ist es auch angesichts der Länge des eingesetzten Fragebogens und der Vielzahl der Fragen (11 Seiten mit 72 Fragen, wenn Statements einzeln gerechnet werden, noch weitaus mehr). In der folgenden Analyse konzentrieren wir uns auf die geschlossenen Fragen, die das Thema Kriminalität bzw. Kriminalitätsfurcht betreffen, und als Vergleichsgröße auf die geschlossenen Fragen, die im weiteren Sinne das Thema Wirtschaft betreffen (einschließlich Arbeitslosigkeit, Rente, Wohnungs- und Umweltfragen). Beide Themen gelten in der Literatur – im Gegensatz etwa zu außenpolitischen Themen – für den einzelnen als subjektiv bedeutsam (Converse 1964). Was bedeuten würde: die Reliabilitäten müßten relativ hoch ausfallen.

3. Reliabilität im Aggregat- und Individualvergleich

Befragungen können auf Seiten der Befragten Reflexionsprozesse auslösen, mal stärker und mal schwächer – je nachdem, wie sehr bereits über das Thema nachgedacht wurde und wie sehr es interessiert (Jagodzinski 1987). Doch in welcher Weise diese Reflexionsprozesse ausgelöst und verarbeitet werden und in welcher Richtung sich die Effekte auswirken, ist bislang kaum erforscht. Auf der einen Seite könnte man meinen, daß durch die Themenstellung der Befragung ein Bezugsrahmen begünstigt wird, der dazu führt, daß *alle* Fragen zum gleichen Thema in akzentuierter Weise wahrgenommen und beurteilt werden (gewissermaßen eine Art „Agenda-Effekt“ analog zu dem Agenda-Effekt in der Medienwirkungsforschung vgl. Schenk 1987). Unter diesen Umständen würde man eine Sensitivierung für die jeweils erfaßten Themenkomplexe erwarten. In unserer Erhebung, bei der die meisten Fragen auf die Kriminalitätsthematik bezogen waren, müßte dies eine Erhöhung der Kriminalitätsfurcht begünstigen. Andererseits könnte man auch mutmaßen, daß die vorgegebene Thematik jeweils im Kontext der eigenen Erfahrungen ihren Bedeutungsgehalt gewinnt. Da in unserer Studie die Kriminalitätserfahrung eine seltene Erfahrung ist – weniger als ein Viertel der Befragten geben an, in den letzten 12 Monaten ein Opfer geworden zu sein –, müßte man eine Anpassung übertriebener Äußerungen (wie sie sich in unseren Daten andeuten) an die Realität und damit einen Rückgang der Furcht erwarten.

Tabelle 1 zeigt die berücksichtigten Items für das Thema Kriminalität mit den entsprechenden Mittelwerten aus dem ersten und zweiten Fragebogen und den Korrelationen (Pearson's r) zwischen den beiden.

Tabelle 1: *Items im Bereich der Kriminalität mit ihren Mittelwerten in den ersten und zweiten Fragebögen und den Korrelationen zwischen den ersten und zweiten Antworten.*

Item	Anzahl der Antwortkategorien	Mittelwert 1	Mittelwert 2	Änderung vom Mittelwert	r
Zufriedenheit mit Schutz vor Kriminalität	5	3,78	3,56	-	.34*
Häufigkeit der Unterhaltung über Kriminalität	5	2,43	2,41	-	.81**
Wichtigkeit der Verbrechensbekämpfung	4	1,30	1,24	-	.65**
Sorge um Wohnungseinbruch	5	2,98	3,00	+	.55**
Sorge, wenn abends allein auf der Straße	5	2,73	2,78	+	.59**
Sorge um Überfall	5	2,93	2,93		.71**
Sorge um die Zunahme der Kriminalität	5	2,18	2,07	-	.70**
Wahrscheinlichkeit, angepöbelt und bedroht zu werden	4	2,64	2,59	-	.64**
Wahrscheinlichkeit, bestohlen zu werden	4	2,51	2,46	-	.87**
Wahrscheinlichkeit, überfallen und beraubt zu werden	4	2,87	2,80	-	.83**
Wahrscheinlichkeit, körperlich angegriffen zu werden	4	2,85	2,88	+	.75**
Wahrscheinlichk., daß in die Wohnung eingebrochen wird	4	2,77	2,78	+	.68**
Wahrscheinlichk., zu sexuellen Handlungen gezwungen zu werden	4	3,38	3,33	-	.57**
Wahrscheinlichkeit, daß das Auto gestohlen wird	4	2,52	2,54	+	.68**
Wahrscheinlichkeit, daß das Auto aufgebrochen wird	4	2,19	2,32	+	.74**
Wahrscheinlichk., daß das Auto mutwillig beschädigt wird	4	2,19	2,25	+	.89**
Sicherheitsgefühl in der Wohngegend abends allein auf der Straße	4	2,47	2,50	+	.62**
Überfallgefahr, wenn man am Tag von der Bank Geld abholt	5	3,20	3,11	-	.74**
Überfallgefahr, wenn man am Tag am Geldautomaten Geld abhebt.	5	3,30	3,28	-	.73**
Überfallgefahr bei Dunkelheit abends und allein zu Fuß in der Innenstadt	5	2,53	2,57	+	.63**
Überfallgefahr abends in Bus/Straßenbahn	5	2,89	2,98	+	.63**
Überfallgefahr bei Dunkelheit in Park- und Grünanlagen	5	2,11	2,16	+	.64**
Überfallgefahr bei Dunkelheit unterwegs von Bus-/Bahnhaltestelle nach Hause	5	2,98	3,02	+	.56**
Gefahr, in der Innenstadt beim Einkaufen bestohlen zu werden	5	2,73	2,86	+	.62**
Gefahr des Wohnungseinbruchs, wenn man auf Reisen ist	5	3,16	3,07	-	.67**

In Tabelle 2 geschieht dasselbe für das Thema Wirtschaft. Bei einem Vergleich zwischen den Mittelwerten der ersten und zweiten Bögen kann man feststellen, daß sich die Mittelwerte in den meisten Fällen nur geringfügig unterscheiden und daß diese Unterschiede nicht systematisch sind. Das heißt, daß in manchen Fällen ein größerer und in manchen Fällen ein kleinerer Mittelwert zu verzeichnen ist. Nur in einem Fall ist die Mittelwertdifferenz auf dem 95 Prozentniveau statistisch signifikant.

Die Tabellen 1 und 2 zeigen weiterhin die Korrelationen zwischen den ersten und zweiten Antworten auf die selben Fragen. Eine schwache Korrelation bedeutet stärkere Abweichungen der Antworten voneinander und eine starke Korrelation eine höhere Übereinstimmung.

Tabelle 2: *Items im Bereich der Wirtschaft mit ihren Mittelwerten in den ersten und zweiten Fragebögen und den Korrelationen zwischen den ersten und zweiten Antworten*

Item	Anzahl der Antwortkategorien	Mittelwert 1	Mittelwert 2	Änderung vom Mittelwert	r
Zufriedenheit mit Angebot an Arbeitsplätzen	5	4,03	4,09	+	.70**
Zufriedenheit mit Wohnungsangebot	5	3,26	3,00	-	.35*
Häufigkeit der Unterhaltung über Atomenergie	5	3,51	3,47	-	.61**
Häufigkeit der Unterhaltung über Sanierung von Wohnungen u. Mietpreise	5	2,07	2,26	+	.52**
Häufigkeit der Unterhaltung über Arbeitslosigkeit	5	2,02	2,00	-	.73**
Beurteilung der eigenen wirtschaftlichen Lage	5	2,60	2,67	+	.83**
Beurteilung der eigenen wirtschaftlichen Lage in einem Jahr	5	2,93	2,91	-	.62**
Wichtigkeit der Schaffung von mehr Arbeitsplätzen	4	1,14	1,09	-	.48**
Wichtigkeit einer stabilen Wirtschaft	4	1,23	1,34	+	.57**
Wichtigkeit des Umweltschutzes	4	1,38	1,49	+	.67**
Wichtigkeit des Sicherns der Renten	4	1,32	1,36	+	.45**
Wichtigkeit der Verbesserung der Lage auf dem Wohnungsmarkt	4	1,58	1,67	+	.53**
Sorge um die Umweltverschmutzung	5	2,41	2,34	-	.52**
Sorge um eigene Arbeitslosigkeit	5	2,61	2,78	+	.82**
Sorge um Kündigung der Wohnung	5	3,60	3,39	-	.71**
Sorge um die Mieterhöhung	5	2,74	2,70	-	.56**
Sorge um die Arbeitslosigkeit in der Familie	5	2,57	2,53	-	.60**

* p<0,05

** p<0,01

Da die selben Personen die selben Fragen zweimal beantwortet haben, wären im Fall einer Konsistenz des Antwortverhaltens sehr starke Korrelationen zu erwarten. Dies trifft in der Regel, aber nicht durchgängig zu. Dabei spielt es keine Rolle, ob die Fragen sich auf das Thema Kriminalität beziehen oder auf die Wirtschaft.

Gleichwohl: auch wenn es gelegentlich zu niedrigen Korrelationen kommt, ist die Gesamtheit der Reliabilitäten zufriedenstellend. Der Durchschnitt der Korrelationen in Tabelle 1 ($r = .67$) unterscheidet sich kaum vom Durchschnitt der 17 Korrelationen in Tabelle 2 ($r = .60$).²

So sehr es auch an einem einheitlichen Trend fehlt und die Reliabilitäten in ihrer Gesamtheit einander gleichen, bleibt als Frage, wie sehr der Wandel auf der Ebene *einzelner* Variablen auf der Aggregatebene mit einem Wandel auf der Individualebene der gleichen Indikatoren einhergeht. Ein Zusammenhang ist denkbar, aber nicht zwingend. Schließlich können Unterschiede in den Antwortmustern auf der Aggregatebene auf der Individualebene sowohl mit Unterschieden als auch Gemeinsamkeiten einhergehen, ebenso wie Gemeinsamkeiten auf der Aggregatebene sowohl mit Unterschieden als auch Gemeinsamkeiten auf der Individualebene parallel gehen können (Reuband 1980).

Um diese Frage zu klären, haben wir zwischen den einzelnen Variablen die absoluten Mittelwertsdifferenzen (also ohne Vorzeichen) mit der Stärke der Korrelationskoeffizienten in Beziehung gesetzt – jeweils getrennt für die Fragen zur Kriminalität und zu den Wirtschaftsfragen. Als Ergebnis zeigt sich dann: es gibt bei beiden Themen keine nennenswerten und statistisch signifikanten Beziehungen.

Klammert man den einzigen Ausreißer aus, der das Gesamtbild des Zusammenhangs bei den Kriminalitätsfragen verfälscht, erhält man über alle Variablen gerechnet im Fall der Kriminalitätsthematik ($N = 23$) eine Korrelation von $r = .05$ und im Fall der Wirtschaftsthematik ($N = 17$) ein nicht signifikantes r von $.15$. Was bedeutet: der Aggregatwandel auf der Ebene einzelner Items wirkt sich nicht auf den Zusammenhang auf der Individualebene aus.

4. Einflüsse auf die Reliabilität

Nun ist es wegen der geringen Fallzahl möglich, daß die beobachteten Abweichungen bzw. Übereinstimmungen zwischen den beiden Bögen bei manchen Variablen in erster Linie zufallsbedingt sind. Um diese Effekte auszuschalten, konzentrieren wir uns in den weiteren Analysen, in denen es um die Einflüsse auf die Reliabilität geht, deshalb nicht auf die Abweichungen bei einzelnen Items, sondern bilden einen Index aus den 25 Items des Themas Kriminalität: den Mittelwert der absoluten Abweichungen der einzelnen Items.

Um die Abweichungen bei den Items mit vier Antwortkategorien mit denen der Items mit fünf Antwortkategorien vergleichbar zu machen, wurden diese vor der Indexbildung mit dem Gewicht $4/3$ multipliziert. Ein ähnlicher Index wurde auch für den Themenbereich Wirtschaft gebildet. Der Index für die Kriminalität weist einen Mittelwert von $.51$ und eine Standardabweichung von $.29$ auf. Die entsprechenden Werte für den Index Wirtschaft lauten $.45$ und $.25$. Somit sind keine gravierenden Unterschiede bei den beiden Verteilungen zu beobachten.

Wir können anhand dieser Indizes nun überprüfen, welche Einflußfaktoren ein inkonsistentes Antwortverhalten der Befragten begünstigen. Dabei interessieren uns sowohl die sozialen Merkmale des Befragten als auch Indikatoren, die als Ausdruck der Teilnahme an der Befragung gedeutet werden können. Bei den sozialen Merkmalen könnte man erwarten, daß Effekte sowohl aus der subjektiven Zentralität der Thematik erwachsen als auch aus dem allgemeinen Grad an Reflexion über gesellschaftliche und politische Themen: je zentraler das Thema und je reflektierter der Befragte ist, desto stabiler die Antworten.

Bei den Indikatoren für Motivation könnte man erwarten, daß die Hochmotivierten weniger zufällig antworten als die wenig Motivierten. So könnte man vermuten, daß bei einer Befragung, die den Befragten gefällt, genauere Antworten geben werden als bei einer Befragung, die nicht gefällt.

Bei den sozialen Merkmalen interessieren uns vor allem die Merkmale Geschlecht, Alter und Bildung. Frauen, schlechter Gebildete und Ältere weisen gewöhnlich eine überproportional hohe Kriminalitätsfurcht auf (vgl. u.a. Boers 1991, Reuband 1999a, b) – wenn dies die subjektive Zentralität des Themas erhöht, müßte sie auch die stabileren Werte über die Zeit erbringen. Andererseits haben schlechter Gebildete bei vielen Themen wenig auskristallisierte Einstellungen, wie Converse (1964, 1970) dokumentiert hat. Schlechter Gebildete sind wohl deshalb auch gewöhnlich stärker anfällig für Effekte der Interviewsituation. Über die sozialen Merkmale hinaus beziehen wir die Variablen „politisches Interesse“ und „Identifikation mit der Stadt“ in die Analyse mit ein.

Bei Converse (1964) hat das politische Interesse eine zentrale Bedeutung für die Auskristallisation von Einstellungen: der Grad der Reflexion über allgemeine politische Themen wird erhöht über die Zeit und auch die Stabilität der Antworten. In unserer Analyse hat vor allem die Identifikation mit der Stadt als Indikator für Befragtenmotivation eine Bedeutung: da die Überschrift des Fragebogens „Umfrage zum Leben in Dresden“ (bzw. Chemnitz oder Leipzig) war und als eine solche auch den Befragten im Anschreiben präsentiert wurde, kann man davon ausgehen, daß diejenigen, die sich mit ihrer Stadt identifizieren, ein größeres Interesse und folglich mehr Konzentration bei der Beantwortung der Fragen aufbrachten. Sie müßten deswegen höhere Reliabilität aufweisen als die übrigen Befragten.

Tabelle 3 dokumentiert die Korrelationen der Indizes mit einigen Befragtenmerkmalen und mit dem zeitlichen Abstand zwischen den ersten und zweiten Bögen.

Dabei zeigt sich: Die aufgeführten Variablen können insgesamt die Inkonsistenzen beim Thema Kriminalität besser erklären als beim Thema Wirtschaft. Und der Effekt der Variablen variiert je nach Thema in seiner Stärke und Richtung. Im einzelnen wird deutlich:

Geschlecht: Es gibt keinen nennenswerten Zusammenhang zwischen den Inkonsistenzen und dem Geschlecht (hier vertreten durch eine Dummy-Variable mit dem Wert 1 für die Frauen und Wert 0 für die Männer).

Alter: Im Gegensatz zu der Annahme von der Bedeutung der subjektiven Zentralität des Themas und einer daraus erwachsenen Stabilität steigen mit zunehmendem Alter die inkonsistenten Antworten beim Thema Kriminalität sehr stark. Beim Thema Wirtschaft ist es eher umgekehrt. Dies bedeutet, daß nicht generell, sondern themenspezifisch gilt, welcher Zusammenhang zwischen der Reliabilität von Fragen und dem Alter besteht.

Bildung: Bei den Befragten mit höherer Bildung ist eine stärkere Übereinstimmung zwischen den Antworten des ersten und des zweiten Bogens zu beobachten, wenn es sich um das Thema Kriminalität handelt. Beim Thema Wirtschaft ist dies nicht der Fall.

Interesse für Politik: Die Befragten mit einem starken Interesse für Politik weisen eine höhere Konsistenz in ihren Antworten auf als die anderen Befragten. Das ist insbesondere beim Thema Wirtschaft der Fall. Womöglich ist das politische Interesse hier deshalb von Bedeutung, weil die Fragen zur allgemeinen wirtschaftlichen Lage von der öffentlichen Thematisierung mitbeeinflusst werden und Personen mit politischen Interessen stärker Informationen dazu aufnehmen.

Tabelle 3: *Korrelationskoeffizienten zwischen Befragten- und Befragungsmerkmalen und der Inkonsistenz der Befragten in den Themenbereichen Kriminalität und Wirtschaft*

	Index Kriminalität	Index Wirtschaft
Alter	.46**	-.16
Geschlecht (weiblich)	.01	.13
Bildung	-.27	.09
Interesse für Politik	-.12	-.29
Gerne leben in der eigenen Stadt	-.12	-.17
Befragung mißfallen I	.19	.07
Befragung mißfallen II	.17	.09
Bereitschaft zum erneuten Interview I	-.50**	-.07
Bereitschaft zum erneuten Interview II	.28	.26
Anzahl der fehlenden Antworten	.37*	-.01
Zeitlicher Abstand zwischen dem ersten und zweiten Bogen	.12	.27

I = 1. Befragung, II = 2. Befragung

* $p < 0,05$

** $p < 0,01$

Gerne leben in eigener Stadt: Wer gern in der eigenen Stadt lebt, weist eine höhere Stabilität der Einstellungen auf. Die negativen Korrelationen dieser Variable mit unseren beiden Indizes stehen im Einklang mit den Erwartungen.

Einstellung zur Befragung: Am Ende des Fragebogens wurden die Befragten gefragt, wie sehr ihnen persönlich die Befragung gefallen hat. Wir haben bei dieser Variable die Antworten aus der ersten *und* aus der zweiten Befragung getrennt be-

rücksichtigt, weil wir davon ausgehen, daß man beim zweiten Ausfüllen des Bogens möglicherweise eine andere Einstellung zu ihm hat (die Korrelation zwischen beiden beträgt $r = .52$). Die positiven Korrelationen auf der Tabelle 3 zeigen, daß beim Mißfallen der Befragung die Inkonsistenzen steigen. Dies ist am ehesten beim Thema Kriminalität der Fall.

Bereitschaft zu einem erneuten Interview: Die Befragten hatten anzugeben, ob sie zu einem späteren Zeitpunkt noch mal an einer Umfrage teilnehmen würden. Wir haben auch hier die Antworten aus den beiden Bögen getrennt berücksichtigt, weil sich diese Bereitschaft nach wiederholter Beantwortung der Fragen möglicherweise ändert. Die Ergebnisse zeigen, daß beim Thema Kriminalität eine größere Inkonsistenz unter den Befragten zu beobachten ist, die nach dem Ausfüllen des ersten Fragebogens nicht bereit sind, noch einmal an der Befragung teilzunehmen. Beim Thema Wirtschaft spielt die im ersten Fragebogen erklärte Bereitschaft praktisch keine Rolle. Die im zweiten Bogen erklärte Bereitschaft zu einer weiteren Befragung geht dagegen mit einer erhöhten Inkonsistenz einher.

Anzahl der fehlenden Antworten: Ein weiterer Index für das Interesse an der Befragung und folglich für die aufgebrachte Aufmerksamkeit ist die Gesamtanzahl der fehlenden Antworten in den beiden Bögen. Dieser Index zeigt einen starken Zusammenhang mit der Inkonsistenz der Antworten, aber nur beim Thema Kriminalität. D.h., daß mindestens in diesem Bereich weniger Antworten keineswegs genauere Antworten bedeuten.

Zeitlicher Abstand zwischen den ersten und zweiten Bögen: Die Diskrepanzen zwischen den Antworten auf die selben Fragen hängen auch mit dem Abstand zwischen dem ersten und zweiten Zeitpunkt zusammen. Das zeigen die positiven Korrelationen dieser Variable mit den beiden Indizes – vor allem im Fall der Frage zur Wirtschaftsthematik. Eine statistische Signifikanz wird allerdings verfehlt, so daß man die Befunde als tendenziell bewerten muß. Bemerkenswert ist, daß sich der Zeitfaktor als weniger bedeutsam erweist als die Motivation zur Teilnahme an der Befragung.

Inwieweit sich die oben genannten bivariaten Zusammenhänge auch in einem multivariaten Modell bewähren, wurde in zwei multiplen Regressionen überprüft. Abhängige Variablen in diesen Regressionen waren die zwei Indizes für die Antwortabweichungen bei den Themen Kriminalität und Wirtschaft. Als erklärende Variablen wurden die in der Tabelle 3 aufgeführten Variablen überprüft. Im Modell Kriminalität blieben nach dem Ausschluß aller nicht signifikanten erklärenden Variablen nur zwei Variablen: „Anzahl der fehlenden Antworten“ und „Bereitschaft zu einem erneuten Interview“ (mit standardisierten Regressionskoeffizienten von $.38^{**}$ und $-.37^{**}$ und einem R^2 von $.38$).

Das Alter, das zunächst in der Korrelationsrechnung im Fall der Kriminalitätsthematik einen bedeutsamen Effekt hatte, übt keinen statistisch signifikanten Effekt mehr aus. Dieses Ergebnis macht deutlich, daß vor allem die Motivation der Befragten eine entscheidende Rolle bei der Genauigkeit der angegebenen Antworten

spielt und daß der Einfluß des Alters primär über die Motivation beim Ausfüllen des Fragebogens erklärt werden kann.

Im Modell Wirtschaft erwies sich keine der erklärenden Variablen als statistisch signifikant, und auch die Regressionsanalyse ändert daran nichts. Möglicherweise ist die Komplexität dieses Themas selbst der Hauptgrund für die ungenauen Antworten der Befragten. Möglicherweise aber auch ist die Wirtschaftsthematik für den einzelnen subjektiv noch bedeutsamer als die Kriminalitätsthematik und die Einstellungen sind entsprechend stärker auskristallisiert.

5. Schlußbemerkungen

Die Angaben zur Kriminalitätsthematik haben sich in der postalischen Befragung als hinreichend reliabel erwiesen. Was bedeutet: die Befragten antworten über die Situationen hinweg konsistent, auch wenn zufallsbedingte Schwankungen im Erleben und der Wahrnehmung miteinwirken. Die Tatsache, daß die Bereitschaft zum erneuten Interview und die Zahl fehlender Antworten im Interview primär die Reliabilität bestimmen, dokumentiert, wie sehr motivations-psychologischen Faktoren bei der Analyse von Reliabilitäten eine zentrale Bedeutung zukommt: Je weniger die Befragten motiviert sind, desto weniger genau beantworten sie die Fragen, und desto stärker wirken Zufälligkeiten der Situation bei der Beantwortung ein.

Forschungsmethodologisch bedeutet dies, daß im Rahmen der Befragungen Anreize gesetzt werden sollten, um die Motivation der Befragten zu erhöhen. Welcher Art diese Anreize sein sollten – auf der Ebene der Fragebogengestaltung, der Fragethemen, des Anschreibens oder anderer Art –, muß im Rahmen weiterer Forschung, u.a. auch über entsprechende Experimente, geklärt werden.

Die Messung der Reliabilitäten hat über die Analyse des Befragtenverhaltens hinaus noch einen weiteren gewichtigen Aspekt. Unter Rekurs auf die Reliabilitätsmessungen, kann man die Stabilität von Einstellungen im längeren Zeitverlauf, so wie sie in Panelstudien normalerweise ermittelt werden, besser einschätzen. In unserem Fall zeigt ein solcher Vergleich, daß eine bemerkenswert hohe Stabilität die Kriminalitätsfurcht über längere Zeiträume hinweg kennzeichnet. Sie liegt z. T. nur unwesentlich niedriger als die Reliabilitäten, die sich im Rahmen der Test-Retest Studie ergaben (vgl. Reuband 1999c).

Anmerkungen:

- 1 Die Untersuchung wurde finanziell gefördert durch das Sächsische Ministerium für Wissenschaft und Kunst (AZ. 4-7541.82-0370/528).
- 2 Reliabilitäten beurteilen wir als zufriedenstellend, wenn sie über .50 – noch besser über .60 – liegen. J. Bortz und N. Döring nennen eine Reliabilität von über .80 als erforderlich, beziehen sich dabei freilich auf Skalen und nicht auf einzelne Items (Bortz und Döring 1995: 84). Skalen weisen immer höhere Reliabilitäten auf, wenn sie sich auf mehrere Items stützen. Auffällig an den Korrelationen ist, daß sich sowohl bei der Kriminalitätsthematik als auch der Wirtschafts-

thematik die „Ausreißer“ nach unten hin auf die allgemeine Zufriedenheit mit der lokalen Infrastruktur und der Effektivität der Infrastruktur beziehen: Zufriedenheit mit dem Schutz der Bürger vor Kriminalität ($r = .34$) und Zufriedenheit mit dem Wohnungsangebot ($r = .35$). Diese Fragen wurden am Anfang des Fragebogens im Rahmen einer Itembatterie gestellt. Inwieweit sich darin Besonderheiten dieser Frage oder deren Dimensionalität widerspiegeln, sei dahingestellt. Die Tatsache, daß die Zufriedenheit mit dem Angebot an Arbeitsplätzen in der gleichen Itembatterie ermittelt wurde und dennoch höhere Korrelationen erreicht wurden, spricht gegen einen Effekt der Fragenkonstruktion und Platzierung.

Literatur:

- Boers, K., 1991: Kriminalitätsfurcht: über den Entstehungszusammenhang und die Folgen eines sozialen Problems. Pfaffenweiler: Centaurus.
- Bortz, J./Döring, N., 1995: Forschungsmethoden und Evaluation für Sozialwissenschaftler. Berlin: Springer.
- Converse, P.E., 1964: The Nature of Belief Systems in Mass Publics. S. 206-261 in: Apter, D.A. (Hrsg.), Ideology and Discontent. New York: The Free Press.
- Converse, P.E., 1970: Attitudes and Non-Attitudes: Continuation of a Dialogue. S. 168-189 in: Tufte, E.R. (Hrsg.), The Quantitative Analysis of Social Problems. Reading, Mass.: Addison Wesley.
- Dillman, D., 1978: Mail and Telephone Surveys. New York: Wiley.
- Jagodzinski, W./Kühnel, S.M./Schmidt, P., 1987: Is there a „Socratic Effect“ in Non-Experimental Panel Studies? Consistency of an Attitude Towards Guest Workers. Sociological Methods and Research 15: 259-302.
- Reuband, K.-H., 1980: Life Histories. Problems and Prospects of Longitudinal Designs. S. 135-163 in: Clubb, J.M./Scheuch, E.K. (Hrsg.), Historical Social Research. The Use of Historical and Process Produced Data. Stuttgart.
- Reuband, K.-H., 1999a: Kriminalitätsfurcht im Alter. Ergebnisse ostdeutscher Studien. S. 209-231 in: Lenz, K./Rudolph, M./Sieckendiek, U. (Hrsg.), Die alternde Gesellschaft. Problemfelder gesellschaftlichen Umgangs mit Altern und Alter. München: Juventa.
- Reuband, K.-H., 1999b: Geschlechtsspezifische Unterschiede in der Kriminalitätsfurcht – eine Folge differentieller Vulnerabilität? Neue Praxis 29: 147-157.
- Reuband, K.-H., 1999c: Kriminalitätsfurcht. Stabilität und Wandel. Neue Kriminalpolitik 11/2: 15-20.
- Reuband, K.-H., 2000a: Der „Standardindikator“ zur Messung der Kriminalitätsfurcht – in „skandalöser Weise“ unspezifisch und in der Praxis dennoch brauchbar? Monatsschrift für Kriminologie und Strafrechtsreform 83: 185-195.
- Reuband, K.-H., 2000b: Telefonische und postalische Bevölkerungsbefragungen in Ostdeutschland. Auswirkungen auf das Antwortverhalten. S. 201-223 in: Hüfken, V. (Hrsg.), Methoden der Telefonumfragen. Opladen: Westdeutscher Verlag.
- Schenk, M., 1987: Medienwirkungsforschung, Tübingen: J. B. Mohr.

Schnell, R./Kreuter, F., 2000: Untersuchungen zur Ursache unterschiedlicher Ergebnisse sehr ähnlicher Viktimisierungsmessungen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 52: 96-117.

Karl-Heinz Reuband, *Heinrich-Heine Universität Düsseldorf,*
Sozialwissenschaftliches Institut, Universitätsstraße, D-40225 Duesseldorf
E-Mail: reuband@phil-fak.uni-duesseldorf.de

Parviz Rastampour, *Martin-Luther-Universität Halle/Wittenberg, Institut für*
Pädagogik, Franckeplatz 1, D-06099 Halle/Saale