

Wie wahrscheinlich ist ‚wahrscheinlich‘?

*Zur subjektiven
Einschätzung und
Kommunikation von
Viktimisierungswahr-
scheinlichkeiten*

How likely is 'likely'?

*Subjective perception and
communication of
victimization probabilities*

*Ivar Krumpal, Heiko Rauhut,
Dorothea Böhr und Elias Naumann*

Zusammenfassung

Was bedeutet ‚wahrscheinlich‘ wenn Be-fragte das Risiko einschätzen, Opfer einer Straftat zu werden? Solche Risiken können einerseits als Gewinne interpretiert und dargestellt werden, indem man hofft, von Straftaten verschont zu bleiben. Andererseits können solche Risiken als Verluste interpretiert werden, indem Respondenten ihre Wahrscheinlichkeit einschätzen, Opfer von Straftaten zu werden. Wir können anhand eines experimentellen Surveys in Leipzig einen Framing-Effekt zeigen: Da ein Verlust schwerer wiegt, werden im Verlust-Frame niedrigere subjektive Viktimisierungswahrscheinlichkeiten angegeben als im Gewinn-Frame. Weiterhin variiert die Zuordnung von prozentualen Wahrscheinlichkeitsangaben zu verbalen Antwortkategorien mit der Häufigkeit eines Ereignisses: Das häufige Ereignis ist ‚von einer Straftat verschont zu bleiben‘, das seltene Ereignis ist ‚einer Straftat zum Opfer fallen‘. Wir können zeigen, dass demselben verbalen Wahrscheinlichkeitslabel (z.B. ‚unwahrscheinlich‘ oder ‚sehr wahrschein-

Abstract

What is the meaning of 'likely', when respondents estimate the risk to become victim of crime? On the one hand, such risks can be presented and interpreted as gains, in terms of being spared of becoming a victim of crime. On the other hand, such risks can be presented and interpreted as losses, in terms of becoming a victim of crime. We demonstrate a framing-effect with data from an experimental survey in Leipzig. Because losses are perceived as more severe, respondents state lower subjective victimization probabilities in the loss-frame, compared to the gain-frame. In addition, the meaning of vague quantifiers varies with the frequency of an event: The frequent event is 'being spared of a crime'; the infrequent event is 'becoming a victim of crime'. We show that the same verbal likelihood labels (e.g. 'unlikely' or 'very unlikely') are more often chosen (in terms of percentages) for frequent than for infrequent events. Finally, with increasing severity of offences, respondents assign for their likeli-

lich') beim häufigen Ereignis höhere prozentuale Wahrscheinlichkeiten zugeordnet werden als beim seltenen Ereignis. Schließlich sinken mit zunehmender Schwere einer Straftat die verbalen Antwortkategorien zugeschriebenen prozentualen Wahrscheinlichkeitswerte. Unsere Ergebnisse zeigen, dass verbale Wahrscheinlichkeitslabel (Vague Quantifiers') nicht wörtlich zu interpretieren sind. Da kategoriale Antwortskalen für verschiedene Viktimisierungen nicht einheitlich verwendet werden, ist ein direkter Vergleich von Anteilswerten über verschiedene Viktimisierungen hinweg als problematisch einzustufen.

hood estimate lower percentages to the same verbal label. Our evidence shows that vague quantifiers cannot be interpreted literally. Because respondents do not use vague quantifiers consistently for different victimizations, it is problematic to compare subjective victimization risks across different victimizations.

1 Einleitung¹

"I thought you'd never come here again," I said.

'Every time I see you, you say the same thing,' she said, laughing. As always, she sat down next to me at the bar and rested both hands on the counter. 'But I did write you a note saying I wouldn't be back for a while, didn't I?'

'For a while is a phrase whose length can't be measured. At least by the person who's waiting,' I said.

'But there must be times when the word's necessary. Situations when that's the only possible word you can use,' she said.

'And probably is a word whose weight is incalculable.'"

"South of the border, west of the sun", Haruki Murakami (2000)

Das Problem, welches der populäre japanische Schriftsteller Haruki Murakami hier in einer poetischen Sprache beschreibt, beschäftigt empirisch arbeitende Sozialwissenschaftler ebenso. Allerdings begnügt sich die Wissenschaft nicht damit, den Wert von ‚for a while‘ oder ‚probably‘ als unmessbar zu bezeichnen. Es gibt bereits einige Versuche, die Bedeutung von solchen unbestimmten verbalen Mengenbegriffen – *Vague Quantifiers* – zu quantifizieren. Simpsons (1944) Aufsatz über Häufig-

1 Wir danken C. Gross, M. Jungbauer-Gans, F. Kreuter, P. Kriwy, M. Trappmann und T. Voss für anregende Diskussionen und wertvolle Hinweise. Wir danken auch den beiden anonymen Gutachtern für ihre konstruktiven Verbesserungsvorschläge. Zudem danken wir T. Köhler und J. Scharf für ihre Unterstützung bei der Datenerhebung. Der Beitrag ist an der Universität Leipzig entstanden.

keitsausdrücke und ihre Bedeutung ist einer der ersten Ansätze. Mittlerweile ist es Konsens, dass nicht jedem verbalen Ausdruck eine feste Zahl zugeordnet werden kann, sondern dass deren Interpretation sowohl intra- als auch interindividuell variiert (vgl. Reyna 1981; Nakao/Axelrod 1983; Wright et al. 1994; Budescu/Wallsten 1995; Kuhn 1997; Karelitz/Budescu 2004).

„One man's ‚rarely‘ is another man's ‚hardly ever“ (Hakel 1969: 533) ist ein anschauliches und wohl auch eines der meistzitierten Beispiele hierfür. Auch für die intraindividuell variierende Verwendung von *Vague Quantifiers* lässt sich ein Beispiel anführen:

„The mean response for ‚sometimes‘ in the shooting in Hollywood westerns context is higher than the mean response for ‚very often‘ in the context of earthquakes.“ (Bradburn 1979: 94).

In der empirischen Sozialforschung ist die Frage nach der Verwendung von *Vague Quantifiers* besonders wichtig, da diese häufig bei Antwortskalen von Standard-items in Fragebögen eingesetzt werden.²

Wir untersuchen Effekte der Frageformulierung sowie der Häufigkeit und Schwere von Delikten auf die subjektive Einschätzung und Kommunikation von Viktimisierungswahrscheinlichkeiten. Zunächst werden wir zeigen, dass bei der Beantwortung von deliktspezifischen Fragen zur eigenen subjektiven Viktimisierungswahrscheinlichkeit mittels prozentualen Wahrscheinlichkeiten Framing-Effekte auftreten. Bereits frühere Studien konnten solche Effekte zeigen (vgl. Tversky/Kahneman 1981). Wir weisen nach, dass derartige Effekte auch im Feld im Rahmen von Surveys auftreten. In unserer zweiten Analyse wird dann der Einfluss des Framings auf die Zuordnung („Mapping“) von *Vague Quantifiers* zu prozentualen Wahrscheinlichkeitswerten untersucht. Es ist festzustellen, dass jedem *Vague Quantifier* im Gewinn-Frame ein höherer prozentualer Wert zugeordnet wird als im Verlust-Frame. Dieser Effekt lässt sich hauptsächlich durch die unterschiedliche Häufigkeit der zu bewertenden Ereignisse erklären („Opfer werden“ ist seltener als „nicht Opfer werden“). Bei einem seltenen

2 „Bitte sagen Sie mir nun für jede der Situationen, die ich Ihnen gleich nenne [z. B. angepöbelt oder bedroht zu werden, zu sexuellen Handlungen gezwungen zu werden, etc.], ob Sie es für *sehr wahrscheinlich*, *wahrscheinlich*, *weniger wahrscheinlich*, oder für *unwahrscheinlich* halten, dass Ihnen persönlich innerhalb der nächsten 12 Monate einmal so etwas passiert?“ (Wohlfahrtssurvey 1998).
 „First, how likely do you think your home is to be burgled in the next year? (...) And how likely do you think you are mugged or robbed in the next year?“ Die möglichen Antwortkategorien waren hierbei: *Very likely*, *fairly likely*, *fairly unlikely* and *very unlikely* (British Crime Survey 2004/2005).

Ereignis wird beispielsweise der verbale Ausdruck ‚oft‘ mit einem kleineren numerischen Wert assoziiert als bei einem sehr häufigen Ereignis (vgl. Bradburn 1979: 94). Allerdings sind auch andere Einflussfaktoren auf das Mapping möglich, so etwa die Schwere eines Deliktes. Wir diskutieren Alternativen zur Verwendung von *Vague Quantifiers*. Wir erörtern, in welchen Forschungskontexten numerische Antwortskalen gegenüber verbalen Antwortskalen vorteilhafter sind.

Unsere Fragestellung ist zudem nicht nur für die Sozialwissenschaften und hier insbesondere für die Fragebogenkonstruktion von Interesse, sondern erhält ihre Relevanz ebenso durch ihre praktische Anwendung in der Medizin. Wie die umfangreiche Literatur aus diesem Bereich zeigt (für einen Überblick vgl. Edwards et al. 2001; Ghosh/Ghosh 2005), wird hier vor allem untersucht, welchen Einfluss der gewählte Frame bei der Kommunikation des Gesundheitsrisikos (beispielsweise die Wahrscheinlichkeit von Nebenwirkungen oder die Wahrscheinlichkeit von Krebs) auf die Entscheidung des Patienten für oder gegen eine Vorsorgeuntersuchung hat (vgl. Gurm/Litaker 2000; Welkenhuysen et al. 2001).³

2 Literaturdiskussion

2.1 *Vague Quantifiers*

Die Verwendung von *Vague Quantifiers* stellt die Vergleichbarkeit von Antworten zwischen verschiedenen Items häufig in Frage (vgl. Kreuter 2000; Schnell/Kreuter 2000; Kreuter 2002: 226-229). Dennoch werden sie, selbst wenn es Alternativen gibt, weiterhin verwendet. Warum aber benutzt man *Vague Quantifiers* und fragt nicht nach dem genauen Wert? Grundsätzlich sind *Vague Quantifiers* die natürliche Antwort auf viele Fragen (vgl. Wallsten et al. 1993). Die Frage, wie oft man in der letzten Woche glücklich war, lässt sich nur schwer mit einer genauen Angabe in Prozent beantworten. Befragte bevorzugen es, Informationen verbal und nicht numerisch zu geben (vgl. Moxey/Sanford 2000), andererseits erhält man Informationen lieber numerisch (vgl. Budescu/Wallsten 1995: 298). Werden Befragte aufgefordert, ihre Antworten quantitativ zu geben, so löst dies in der Regel zunächst

3 Je nach dem, ob man das an sich gleiche Risiko von Nebenwirkungen im Gewinn-Frame oder Verlust-Frame kommuniziert, entscheidet sich ein Patient eher gegen bzw. für eine Vorsorgeuntersuchung.

ein Nachdenken aus, nur die wenigsten verweigern jedoch eine Antwort (vgl. Bradburn 1979: 95). Präzise quantitative Schätzungen erfordern eine hohe kognitive Leistung (vgl. Peters et al. 2006: 46). So geht in der Fragesituation viel Zeit und Konzentration für die Antwortfindung verloren. Gleichzeitig besteht das Risiko, dass die genaue Zahl gar nicht vorhanden ist („Non-Attitude“), sondern erst vom Befragten situativ generiert wird (vgl. Conrad et al. 1998: 363).

Da numerische Angaben häufig eine Genauigkeit suggerieren, die für den Befragten so nicht vorliegt, erscheint die Verwendung von *Vague Quantifiers* adäquater – wenn man nur wüsste, was sie bedeuten. Wenn bei der Auswertung bekannt wäre, welchem *Vague Quantifier* welcher exakte numerische Wert zugeordnet ist, könnte man weiterhin die kognitiv einfacher zu verstehenden *Vague Quantifier* in Befragungen nutzen. Das sogenannte ‚Mapping‘ realisiert diese Zuordnung von exakten numerischen Werten zu *Vague Quantifiers* (vgl. Wright et al. 1994: 481).

Theoretische Grundlage für das Mapping ist die Annahme von ‚Membership Functions‘ (Hammerton 1976). Es wird davon ausgegangen, dass Wahrscheinlichkeitsausdrücke ungenaue Konzepte sind, und dass verschiedene numerische Wahrscheinlichkeiten zwischen 0 und 1 diese Ausdrücke repräsentieren. Die Membership Function eines verbalen Ausdrucks ordnet jedem Wert zwischen 0 und 1 eine Zahl zu, die angibt, wie stark diese Wahrscheinlichkeit mit der verbal ausgedrückten Häufigkeit korrespondiert.

„Probability phrases are vague concepts and (...) different numerical probabilities in the [0, 1] range are represented to various degrees in these concepts. (...) The membership function of any given phrase assigns a number to each value on the probability line [0, 1] that represents its degree of membership in the concept defined by the phrase.“ (Karelitz/Budescu 2004: 27).

Erste Studien nahmen an, dass jeder Ausdruck nur eine Membership Function hat. Es wurde versucht, jedem *Vague Quantifier* genau einen numerischen Wert zuzuordnen (vgl. Hammerton 1976). Relativ schnell wurde aber klar, dass die Membership Functions und damit auch das Mapping von vielen Faktoren beeinflusst werden.

Vor allem psychologische Studien heben die Wichtigkeit des Kontextes hervor. Goocher (1965) betont, dass es eine Rolle spielt, ob man das Ereignis oder das Objekt mag oder nicht. Je mehr die Respondenten ein Ereignis oder ein Objekt mochten, desto kleinere Häufigkeitsausdrücke wurden verwendet. Daran anknüpfend zeigen Pepper/Prytulak (1974) sowie Moxey/Sanford (1993, 2000) den Einfluss der erwarteten Häufigkeit:

„In particular, if an event has a high base-rate expectation, such as people enjoying parties, then the values assigned to (say) *many* in *many* people enjoyed the party is higher than it is for a low baserate expectation (as in *many* of the doctors in the hospital were female).“ (Moxey/Sanford 2000: 241).

Schwarz et al. (1985) verweisen auf den Einfluss der Skalierung der Antwortvorgaben. Wenn man die niedrigeren Kategorien einer Antwortskala ausdifferenziert, geben die Befragten einen niedrigeren Fernsehkonsum an. Im höheren Bereich ausdifferenzierte Antwortkategorien ergeben dagegen höhere Häufigkeitsangaben des abgefragten Verhaltens.⁴ Hörmann (1983) und Newstead (2000) wiederum beschäftigen sich mit der Variation, die durch die veränderte Größe von Objekten ausgelöst wird.

„[...] *a few* people in front of a building meant more than *a few* people standing in front of a hut“ (Newstead 2000: 244).

Neben individuellen Variationen der Assoziationen von *Vague Quantifier* mit numerischen Angaben bzw. neben Manipulationen, die alle Befragten im gleichen Maße betreffen, ist es für die sozialwissenschaftliche Forschung problematischer, wenn Unterschiede im Verständnis von *Vague Quantifiers* gruppenspezifisch auftreten. Festgestellte Gruppenunterschiede beruhen dann möglicherweise nicht auf tatsächlichen Unterschieden, sondern auf einer unterschiedlichen Interpretation von *Vague Quantifiers*. Eine Studie von Schaeffer (1991) verdeutlicht dies: Sie befragte schwarze und weiße Teilnehmer nach dem Ausmaß ihrer Langeweile. Wurden numerische Werte verwendet, waren keine Unterschiede festzustellen. Bei der Verwendung von *Vague Quantifiers* gaben schwarze Teilnehmer allerdings an, häufiger gelangweilt zu sein. Nakao/Axelrod (1983) weisen Gruppenunterschiede zwischen Ärzten und Laien im Gebrauch von verbalen Häufigkeitsausdrücken wie ‚infrequent‘ oder ‚not infrequently‘ nach. King et al. (2004) zeigen, dass die Verwendung von Antwortskalen interkulturell variiert.

Wänke (2002) zeigt, dass es wichtig ist, mit welcher Referenzgruppe man sich vergleicht – die gesamte Bevölkerung oder die eigene soziale Gruppe. So schätzen Studenten, die sich mit der gesamten Bevölkerung vergleichen, ihre Kino-

4 Neben der Frage nach dem Fernsehkonsum konnten vergleichbare Effekte auch bei der Abfrage anderer Verhaltensweisen, wie etwa sexuelles Verhalten oder Konsumverhalten, festgestellt werden (vgl. Schwarz/Scheuring 1988; Menon et al. 1995). Ähnliche Effekte konnten auch bei der Frage nach Häufigkeiten körperlicher Symptome bei Patienten beobachtet werden (vgl. Schwarz/Scheuring 1992).

frequenz höher ein als Studenten, die sich mit anderen Studenten vergleichen. Auch spielt es eine Rolle welches Referenzverhalten man wählt:

„For example, when judging how often one goes to the movies, the response may depend on whether one compares this behaviour to the frequency one attends poetry readings or watches TV.“ (Wänke 2002: 302).

Wright et al. (1994) fassen die unterschiedlichen Faktoren, welche das Mapping beeinflussen, zu zwei Kategorien zusammen. Zum einen hängt das Mapping von Informationen ab, die der eigenen Erfahrung entspringen. Sollen den unbestimmten Adjektiven ‚jung‘, ‚alt‘, ‚sehr alt‘ ... numerische Altersangaben zugeordnet werden, dann spielt das eigene Alter eine Rolle. Diese Annahme wird als ‚Self-Information-Hypothese‘ bezeichnet. Zum anderen wird das Mapping durch Eigenschaften (Verhalten, Sprachgebrauch etc.) der sozialen Gruppe bestimmt, zu der man gehört. Ein Beispiel, das einen dieser Aspekte verdeutlicht, findet sich ebenfalls bei Wright et al. (1994): In zwei isolierten Dörfern werden die Bewohner befragt, wie groß sie seien. Als Antwortmöglichkeiten sind *Vague Quantifiers* vorgegeben. Angenommen, in dem einen Dorf sind alle Personen 1,80 m groß und im anderen Dorf sind alle Personen 1,90 m groß. Soll der Befragte nun seine Größe einschätzen, so denkt er an die anderen in seinem Dorf und kommt zu der Annahme, er sei normal groß. Entsprechend ordnet er sich in der Mitte der vorgegebenen Skala ein. Bei der Erfassung der Größe mittels *Vague Quantifiers* kommt man in diesem Fall also zu dem Ergebnis, dass die Bewohner beider Dörfer genau gleich groß sind – was aber nicht der Fall ist. Die Annahme, dass die Zugehörigkeit zu einer sozialen Gruppe das Mapping beeinflusst, wird als ‚Group-Norm-Hypothese‘ bezeichnet. Wright et al. (1994) bestätigen in ihrer Studie zum Fernsehkonsum⁵ die Gültigkeit beider Hypothesen, regen aber an, dass weitere Untersuchungen zu Gruppenunterschieden nötig seien, um ihre Schlussfolgerungen zu untermauern.

2.2 Effekte des Framings

Die subjektive Einschätzung von Wahrscheinlichkeiten hängt davon ab, wie solche Risiken präsentiert werden. Tversky/Kahneman (1981) beschreiben in ihrer mittlerweile klassischen Studie zum Framing von Entscheidungen zwei unterschiedliche

5 Wenn erst nach dem Fernsehkonsum des ‚typisch Anderen‘ gefragt wird, geben Respondenten in der folgenden Frage nach dem eigenen Konsum mehr Stunden an, als wenn sie zuerst nach ihrem eigenen Konsum gefragt werden.

Szenarien der Seuchenbekämpfung, ein sicheres und eine risikoreiches. Einmal wird das Ergebnis mit der Anzahl der Überlebenden beschrieben, das andere Mal mit der Anzahl derer, die sterben. Obwohl das Entscheidungsproblem in beiden Formulierungen den gleichen statistischen Erwartungswert hat, und somit formal äquivalent ist, entscheiden sich die Befragten für die risikoreiche Variante, wenn sie mit dem Verlust-Frame (Anzahl der Toten) konfrontiert werden, und für die sichere Variante, wenn sie mit dem Gewinn-Frame konfrontiert werden.

Laut Prospect Theorie (Kahneman/Tversky 1979) erzeugt die Darstellung eines Entscheidungsproblems unter Unsicherheit in Form von Gewinnen im Vergleich zu Verlusten eine Verschiebung des Referenzpunktes, von dem aus die Befragten ihre Evaluation durchführen. So ist der Referenzpunkt ein anderer, wenn man bei einem Risiko einerseits den drohenden Verlust in den Vordergrund stellt oder andererseits den Gewinn, von diesem Risiko verschont zu bleiben. Die Annahme der Prospect Theorie ist, dass erstens die Bewertungsfunktion für Gewinne konkav und für Verluste konvex verläuft und zweitens, dass der Anstieg der Bewertungsfunktion für Verluste steiler ist als für Gewinne. Diese Eigenschaften führen dazu, dass die gleiche Wahrscheinlichkeit des Eintretens eines bestimmten Ereignisses im Verlust-Frame stärker psychologisch wirksam ist als im Gewinn-Frame. In anderen Worten: zu 5% einen Verlust zu erleiden wiegt stärker, als zu 95% von diesem Verlust verschont zu bleiben.

McNeil et al. (1982) konnten diesen Effekt im klinischen Bereich bestätigen. Patienten hatten zu entscheiden, ob sie eine sicherere Strahlentherapie oder eine risikoreichere Operation bevorzugen. Einmal wurden die Risiken beider Behandlungen in Form von Sterberisiken, einmal in Form von Überlebenswahrscheinlichkeiten dargestellt. Wie erwartet zeigte es sich, dass die Darstellung in Form von Sterberisiken weniger Befragte dazu bewegte, sich für die risikoreichere Operation zu entscheiden, als dies bei der Darstellung in Form von Überlebenswahrscheinlichkeiten der Fall war. Für eine Übersicht dieser Darstellungseffekte in der Medizin siehe Marteau (1989), Banks et al. (1995) und Edwards et al. (2001).

In eine ähnliche Richtung für den Bereich von Konsumententscheidungen gehen die Überlegungen von Moxey/Sanford (2000). Den Befragten werden zwei logisch äquivalente Möglichkeiten vorgegeben, zwischen denen sie sich entscheiden sollen, beispielsweise einem Joghurt, der 95% fettfrei ist und einem, der 5% Fett enthält. Die Entscheidung fällt meist auf den 95% fettfreien Joghurt. Als entscheidendes Kriterium für diese Entscheidung betonen sie die Perspektive, die durch die Formulierung gesetzt wird – das eine Mal wird der positive Aspekt (fettfrei) betont, das andere Mal wird die Aufmerksamkeit auf den negativen Aspekt

gelenkt. Im ‚Joghurtproblem‘ ist die Bevorzugung des positiven Frames (95% fettfrei) gegenüber dem negativen Frame (5% Fett) auf eine innerlich höhere Bewertung im Gewinn-Frame zurückzuführen. Man müsste folglich im Gewinn-Frame den Fettgehalt des Joghurts heraufsetzen (z. B. 90% fettfrei), um eine subjektive Äquivalenz mit der Darstellung im Verlust-Frame (5% Fett) zu erhalten. In diesem Fall wäre die Versuchsperson indifferent zwischen den beiden Formulierungen.

Das Ergebnis, dass logisch äquivalente Ausdrücke unterschiedlich bewertet werden, wird ebenso von einer früheren Studie von Reyna (1981) bestätigt. Sie beschäftigt sich mit Modaladjektiven (wie ‚wahrscheinlich‘ oder ‚möglicherweise‘) und der Wirkung, die Verneinungen haben. Dabei stellt sie fest, dass Verneinungen nicht automatisch zu komplementären numerischen Schätzungen führen.

3 Subjektive Einschätzung von Viktimisierungswahrscheinlichkeiten

3.1 Hypothesen

Wenn wir die bisherigen Analysen auf subjektive Viktimisierungswahrscheinlichkeiten übertragen, ist Folgendes zu erwarten: Es wird vorgezogen ‚zu 95% von einem Verbrechen verschont zu bleiben‘ anstelle ‚zu 5% Opfer zu werden‘. Der Optimismus im Gewinn-Frame würde dabei auf einer, im Vergleich zum Verlust-Frame, niedrigeren subjektiven Viktimisierungswahrscheinlichkeit beruhen. Folglich wird erwartet, dass die Wahrscheinlichkeit, verschont zu bleiben (Gewinn-Frame) herabgesetzt wird, damit sie mit der Wahrscheinlichkeit, Opfer zu werden, subjektiv korrespondiert. Ein Beispiel wäre, dass ‚zu 90% verschont bleiben‘ als äquivalent aufgefasst wird wie ‚zu 5% Opfer werden‘. Unsere erste Hypothese lautet damit wie folgt: Im Gewinn-Frame (‚verschont werden‘) erwarten wir höhere subjektive Viktimisierungswahrscheinlichkeiten als im Verlust-Frame (‚Opfer werden‘).

3.2 Methode

Im Rahmen eines experimentellen Surveys zum Thema ‚Kriminalitätsfurcht‘, befragten wir im August 2006 in Leipzig 95 Taxifahrerinnen bzw. Taxifahrer und 96 Kellnerinnen bzw. Kellner (N=191). Erhoben wurden u. a. Kriminalitätsfurcht, Opferwerdung, subjektive Kosten einer Viktimisierung und Wehrhaftigkeit. Außerdem wurden die Befragten gebeten, ihre Wahrscheinlichkeit einzuschätzen, dass

sie selbst Opfer eines Verbrechens werden. Insgesamt befragten wir zu fünf Delikten, die wir in kurzen Szenarien vorstellten: Verkehrsunfall, Zechprellerei, Diebstahl, Raub, Mord. Wir wählten zwei Gruppen von Personen, die ihren Beruf in der Öffentlichkeit ausüben und somit einem höheren Viktimisierungsrisiko ausgesetzt sind. Die Ziehung der Probanden erfolgte mittels einer zweistufigen Klumpenstichprobe, wobei Taxistände bzw. Gastronomiebetriebe jeweils die Klumpen bildeten, in denen in einem zweiten Schritt eine oder mehrere Personen befragt wurden: Zur Auswahl der Taxifahrerinnen und Taxifahrer verwendeten wir eine Liste aller Taxistände der Taxigenossenschaft, aus der wir zunächst zufällig Taxistände zogen. An den einzelnen Ständen fingten wir jeweils mit der Fahrerin bzw. dem Fahrer des zweiten Taxis in der Reihe an und befragten weiter bei jedem zweiten Wagen. Die Anzahl der realisierten Interviews pro Taxistand variierte, je nachdem, wie frequentiert der Taxistand war. Zum anderen lag uns eine Liste aller eingetragenen Gastronomiebetriebe der IHK Leipzig vor, aus der wir eine Zufallsstichprobe zogen. In den einzelnen Betrieben gingen wir nach der Geburtstagsauswahl (last-birthday-method') vor. Die Analyse beruht also auf einer direkten Erhebung aus dem Feld, die jederzeit reproduziert werden kann.

Es wurden von vier Interviewern mündliche Face-to-face Interviews mit einem standardisierten Fragebogen durchgeführt. Dabei gab es zwei Versionen mit unterschiedlichen Frames. Die Zufallsallokation zu den beiden Fragebogenversionen erfolgte mittels Zufallszahlen. Die Randomisierung fand sowohl innerhalb der beiden Berufsgruppen als auch innerhalb der Interviewer statt. Jeweils die Hälfte der TaxifahrerInnen und KellnerInnen fragten wir nach ihrer Schätzung, Opfer eines Verbrechens zu werden (Verlust-Frame), die andere Hälfte beider Berufsgruppen, für wie wahrscheinlich sie es halten, von einem bestimmten Verbrechen verschont zu bleiben (Gewinn-Frame). So wurden 41 KellnerInnen und 49 TaxifahrerInnen im Verlust-Frame und 55 KellnerInnen und 46 TaxifahrerInnen im Gewinn-Frame befragt. Der Fragebogen beinhaltete außerdem eine Kontrollfrage zum Delikt Raub (intraindividuelle Variation des Frames). So bekam die eine Hälfte der Befragten die Frage nach der Viktimisierungswahrscheinlichkeit zweimal im Verlust-Frame gestellt, die andere Hälfte der Befragten sollte erst im Verlust- und dann im Gewinn-Frame antworten. Die Befragten sollten ihre Einschätzung zuerst auf einer verbalen Skala von *Vague Quantifiers* und direkt anschließend noch einmal als numerische Prozentangabe angeben:

Tabelle 1 Die Fragen des Fragebogens

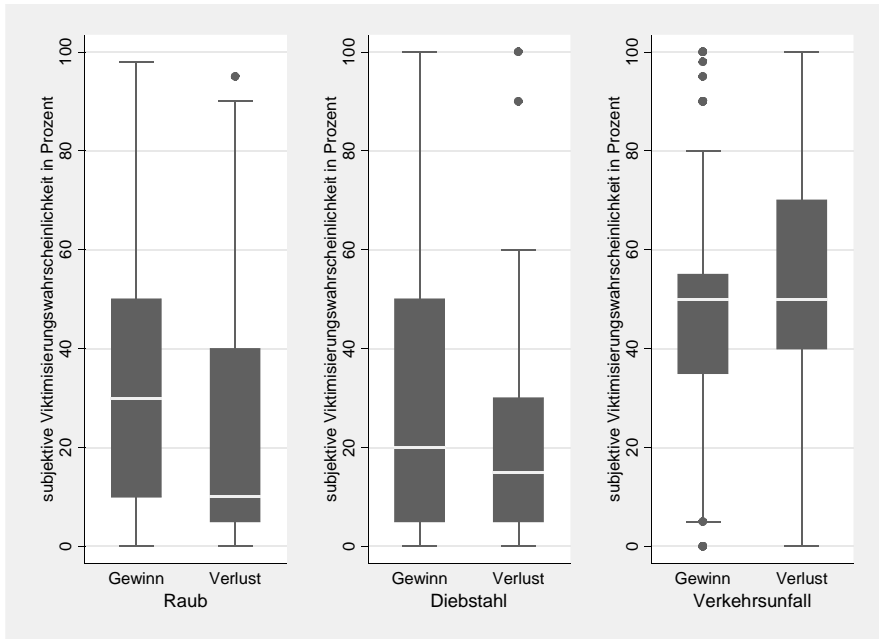
Frage 1:	<p>Ich möchte als Erstes gern über Ihre persönliche Gefährdung sprechen. Dazu nenne ich Ihnen verschiedene Situationen und würde Sie bitten, mir anhand dieser Karte zu sagen, für wie wahrscheinlich Sie es halten, dass Sie in den nächsten 12 Monaten in eine solche Situation geraten.</p> <p>Für wie wahrscheinlich halten Sie es, dass in den nächsten 12 Monaten ...</p> <p>... ein Kunde abhaut ohne zu zahlen?</p> <p>... Ihnen jemand Gewalt androht, um an Ihr Geld oder Ihre Wertgegenstände zu kommen?</p> <p>... Sie jemand umbringt?</p>
<hr/>	
Frage 9:	
<i>Version 1</i>	<p>Für wie wahrscheinlich halten Sie es, in den nächsten 12 Monaten von folgender Situation ... verschont zu bleiben? (...) Für wie wahrscheinlich halten Sie es, dass Ihnen folgendes <u>nicht</u> passiert:</p>
<i>Version 2</i>	<p>Für wie wahrscheinlich halten Sie es, in den nächsten 12 Monaten in folgende Situation zu geraten? (...) Für wie wahrscheinlich halten Sie es, dass Ihnen folgendes passiert:</p> <p>... Ein Kunde stiehlt Ihr Portemonnaie und verschwindet.</p> <p>... Jemand droht Ihnen Gewalt an, um an Ihr Geld oder Ihre Wertgegenstände zu kommen.</p> <p>... Ein anderer Verkehrsteilnehmer nimmt Ihnen die Vorfahrt und es kommt zu einem Zusammenstoß.</p>
<hr/>	
Antwort:	<p>1 – auf keinen Fall, 2 – sehr unwahrscheinlich, 3 – unwahrscheinlich, 4 – teils/teils, 5 – wahrscheinlich, 6 – sehr wahrscheinlich</p> <p>Geben Sie dazu bitte auch noch die Wahrscheinlichkeit in Prozent an.</p>

Die Fragen wurden nicht direkt nacheinander gestellt, es lagen etwa fünf Minuten Zeit zwischen der Beantwortung.

3.3 Ergebnisse

Zunächst betrachten wir nur die Ergebnisse auf der prozentualen Skala und überprüfen den Effekt des Framings auf die subjektive Viktimisierungswahrscheinlichkeit in der numerischen Einschätzung. Mit Boxplots vergleichen wir die Verteilungen pro Frame für die Delikte Diebstahl, Raub und Verkehrsunfall.

Abbildung 1 Boxplots der Verteilungen der subjektiven Viktimisierungswahrscheinlichkeit in beiden Frames pro Delikt. Im Gewinn-Frame wird die Gegenwahrscheinlichkeit dargestellt.



Die Daten zeigen, dass bei den Delikten ‚Raub‘ und ‚Diebstahl‘ wie erwartet eine höhere subjektive Viktimisierungswahrscheinlichkeit im Gewinn-Frame angegeben wurde. So liegen bei beiden Delikten der Median der empirischen Verteilungen im Gewinn-Frame über dem Median des Verlust-Frames. Im Delikt ‚Verkehrsunfall‘ liegen die beiden Mediane dagegen auf gleicher Höhe. Für einen besseren Vergleich der Wahrscheinlichkeiten betrachten wir die Mittelwerte (vgl. Tabelle 2). Die mittleren subjektiven Viktimisierungswahrscheinlichkeiten sind bei den Delikten ‚Raub‘ und ‚Diebstahl‘ in den beiden Frames unterschiedlich. Die Wahrscheinlichkeit verschont zu bleiben wird, wie theoretisch vermutet, herabgesetzt. Die subjektive Viktimisierungswahrscheinlichkeit im Gewinn-Frame wird somit höher eingeschätzt. So beträgt beim Delikt ‚Raub‘ der Mittelwert der eigenen subjektiven Wahrscheinlichkeit 35,5% im Gewinn-Frame (‚Verschont-Bleiben‘). Demgegenüber wird im Verlust-Frame (‚Opfer-Werden‘) eine geringere mittlere subjektive Viktimisierungswahrscheinlichkeit von 21,9% angegeben ($p < 0,01$). Auch beim Delikt ‚Diebstahl‘ ist die Differenz signifikant in Richtung unserer Hypothese ($p < 0,01$). In der folgenden Tabelle sind alle Mittelwerte aufgelistet.

Tabelle 2 Arithmetische Mittel der subjektiven Viktimisierungswahrscheinlichkeit (in Klammern: N)

	Raub	Diebstahl	Verkehrsunfall
Verlust-Frame	21,9 (87)	20,6 (87)	51,4 (85)
Gewinn-Frame (100-p)	35,5 (99)	31,0 (99)	48,8 (96)
p-Wert (einseitiger t-Test)	0,0003	0,0042	0,7560

Im Delikt ‚Verkehrsunfall‘ liegen dagegen die beiden Wahrscheinlichkeiten sehr nahe beieinander (51,4% im Verlust-Frame versus 48,8% im Gewinn-Frame).⁶ Wir können mit den beiden Delikten ‚Raub‘ und ‚Diebstahl‘ im Sinne unserer Hypothese zeigen, dass es einen Einfluss des Framings in der Fragestellung auf die numerischen Angaben in der Antwort gibt.

4 Kommunikation von Viktimisierungswahrscheinlichkeiten

4.1 Hypothesen

Während in der ersten Analyse gezeigt wurde, dass die subjektive Einschätzung von Viktimisierungswahrscheinlichkeiten durch die Frageformulierung beeinflusst wird, untersuchen wir in einer weiteren Analyse, inwiefern subjektive Viktimisierungswahrscheinlichkeiten je nach Situation unterschiedlich kommuniziert werden. Eine unterschiedliche Kommunikation von Risiken drückt sich darin aus, dass die Ausdrücke ‚wahrscheinlich‘ oder ‚unwahrscheinlich‘ in verschiedenen Situationen unterschiedlich verwendet werden. Hierzu analysieren wir die Zuordnung dieser *Vague Quantifiers* zu prozentualen Wahrscheinlichkeitswerten.⁷ Dieses Vorgehen wird als ‚Mapping‘ bezeichnet (vgl. Wright et al. 1994). Wir demonstrieren mit zwei Mapping-Analysen, dass verbale Wahrscheinlichkeitsausdrücke zu Viktimisierungsrisiken je nach Situation eine unterschiedliche Bedeutung haben. *Erstens* vergleichen wir

6 Dies kann mitunter darauf zurückzuführen sein, dass es sich bei einem Verkehrsunfall um ein relativ häufiges Ereignis handelt.

7 Einen Vergleich von verbalen und prozentualen Antwortskalen zu subjektiven Entdeckungswahrscheinlichkeiten bei selbstberichteter Delinquenz anhand von ALLBUS-Indikatoren der Jahre 1990 und 2000 führte unlängst Reuband (2002) durch.

die Interpretation einer gegebenen verbalen Viktimisierungswahrscheinlichkeit (z. B. ‚wahrscheinlich‘) im Gewinn-Frame mit derjenigen im Verlust-Frame. Hierbei wird über das Framing die Häufigkeit des zu bewertenden Ereignisses randomisiert: Das Ereignis ‚Opfer werden‘ ist seltener als das Ereignis ‚nicht Opfer werden‘. Es wird angenommen, dass bei seltenen Ereignissen (z. B. ‚Raub‘) eine gegebene verbale Antwortkategorie mit kleineren numerischen Wahrscheinlichkeiten in Verbindung gebracht wird als bei häufigen Ereignissen (z. B. ‚nicht Raub‘).⁸ Die zu testende Hypothese lautet somit wie folgt: Im Gewinn-Frame („nicht Raub“) werden demselben verbalen Antwortlabel (z. B. ‚sehr wahrscheinlich‘) höhere prozentuale Wahrscheinlichkeiten zugeordnet als im Verlust-Frame („Raub“). *Zweitens* analysieren wir, ob mit zunehmender Schwere eines Deliktes, die prozentualen Wahrscheinlichkeiten, die einem gegebenen verbalen Antwortlabel zugeschrieben werden, sinken. So ist etwa zu erwarten, dass Befragte mit einer ‚sehr wahrscheinlichen‘ Viktimisierung beim Delikt ‚Mord‘ ein geringeres Risiko kommunizieren möchten als bei einer ‚sehr wahrscheinlichen‘ Viktimisierung beim Delikt ‚Zechprellerei‘.⁹

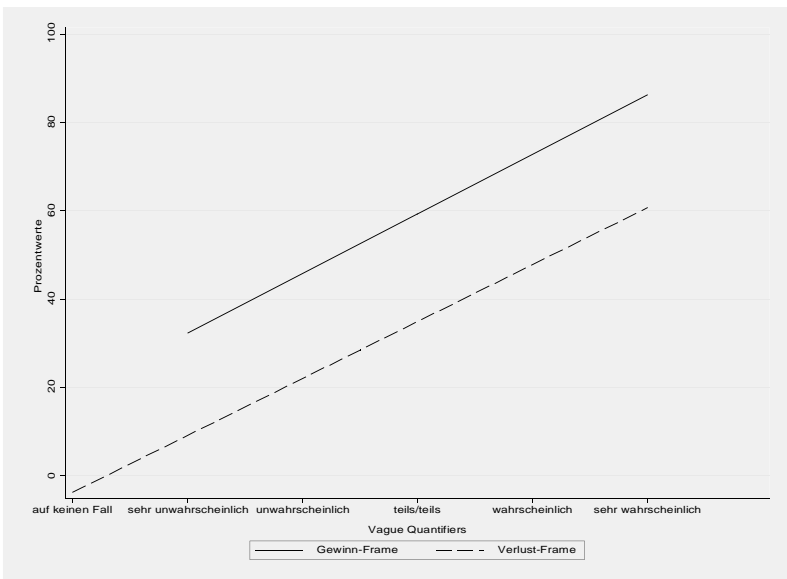
- 8 „In short ‚often‘ for an improbable event is less than often for a highly probable event.“ (Bradburn 1979: 94). Der Häufigkeitseffekt kann anhand der Konversationslogik von Grice (1993) begründet werden. Hierbei wird eine Orientierung des Befragten am Relevanzprinzip und an der Kategorie der Relation angenommen. Wenn der Befragte die Wahrscheinlichkeit einschätzt, im nächsten Jahr Opfer eines Raubs zu werden, und die entsprechende Frage in Form einer absoluten Wahrscheinlichkeit beantwortet, ist die vermutete Antwort ‚sehr unwahrscheinlich‘. Da andere Befragte eine ähnliche Antwort geben werden, wäre diese Information nicht relevant. Den Überlegungen von Grice folgend vermutet der Befragte vielmehr, dass die Wahrscheinlichkeit in Relation zu allen anderen Befragten interessiert. Deshalb wird er die Antwortskala an der vermuteten durchschnittlichen Wahrscheinlichkeit verankern und sich entsprechend einordnen. Für eine detaillierte Beschreibung der „anchoring and adjustment heuristic“ vgl. Peters et al. 2006. Dieser Antwortprozess ist zunächst unabhängig von der Häufigkeit und findet Anwendung sowohl für numerische, als auch für verbale Skalen. Bei der Verwendung der verbalen Skala stehen jedoch qualitative Überlegungen im Vordergrund (vgl. Zimmer 1983; Budescu/Wallsten 1995: 303). Sowohl beim häufigen („nicht Raub“) als auch beim seltenen Ereignis („Raub“) ist eine relevante und relationale Antwort möglich, wenn die eigene Wahrscheinlichkeit in Relation zur vermuteten durchschnittlichen Wahrscheinlichkeit der anderen evaluiert, und diese durchschnittliche Wahrscheinlichkeit mit der Mittelposition der Antwortskala als Anker gleichgesetzt wird. Bei der numerischen Skala stehen hingegen in erster Linie quantitative Überlegungen im Vordergrund. So wird bei häufigen Ereignissen („nicht Raub“) der Anker höher gesetzt als bei seltenen Ereignissen („Raub“). Dies begründet den Häufigkeitseffekt beim Mapping.
- 9 Die beiden Dimensionen ‚Häufigkeit‘ und ‚Schwere‘ eines Deliktes konnten hier nicht unabhängig voneinander manipuliert werden. Um eine mögliche Konfundierung in zukünftigen Studien auszuschließen, sollten Anwendungen entwickelt werden, bei denen dies möglich ist. Eine denkbare Möglichkeit in der Medizin wäre, bei der Untersuchung unterschiedlicher Krebserkrankungen ‚Häufigkeit‘ (Erkrankungsrisiko) und ‚Schwere‘ (Sterberisiko) unabhängig voneinander zu manipulieren. Zu einer vertieften Diskussion dieses Sachverhaltes im medizinischen Kontext vgl. Weber/Hilton (1990); Merz et al. (1991); Sutherland et al. (1991); Budescu/Wallsten (1995: 294f.).

4.2 Einfluss der Häufigkeit eines Ereignisses

Wir rechnen sechs einfache lineare OLS-Regressionen. Für die Delikte ‚Raub‘, ‚Diebstahl‘ und ‚Verkehrsunfall‘ (vgl. Tabelle 1, Frage 9) regressieren wir die Wahrscheinlichkeitsangaben der verbalen Antwortskala auf die zugehörigen Prozentangaben. Wir berechnen jeweils drei Regressionen für den Gewinn-Frame und drei Regressionen für den Verlust-Frame. Die daraus resultierenden Response Functions werden für jedes Delikt graphisch abgebildet (vgl. Wright et al. 1994).¹⁰ Auf der X-Achse werden die Werte der verbalen Antwortskala und auf der Y-Achse die Werte der prozentualen Wahrscheinlichkeitsangaben abgetragen.

Unsere Vorhersage, wonach den *Vague Quantifiers* im Gewinn-Frame (also dem häufigeren Ereignis des Nicht-Eintretens einer Opferwerdung) höhere prozentuale Wahrscheinlichkeiten zugeschrieben werden, lässt sich bestätigen (vgl. Abbildungen 2-4).

Abbildung 2 Inter-personelles Framing, Delikt: Raub (Frage 9)



10 Wright et al. (1994) berechnen in ihrer Studie zum Mapping von *Vague Quantifiers* ebenfalls OLS-Regressionen und interpretieren die daraus gebildeten Regressionsgeraden als ‚Response Functions‘.

Abbildung 3 Inter-personelles Framing, Delikt: Diebstahl (Frage 9)

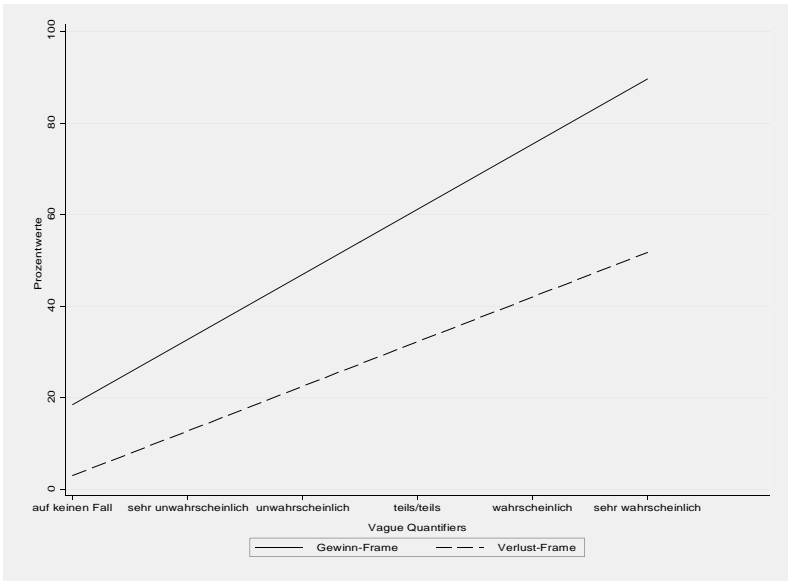
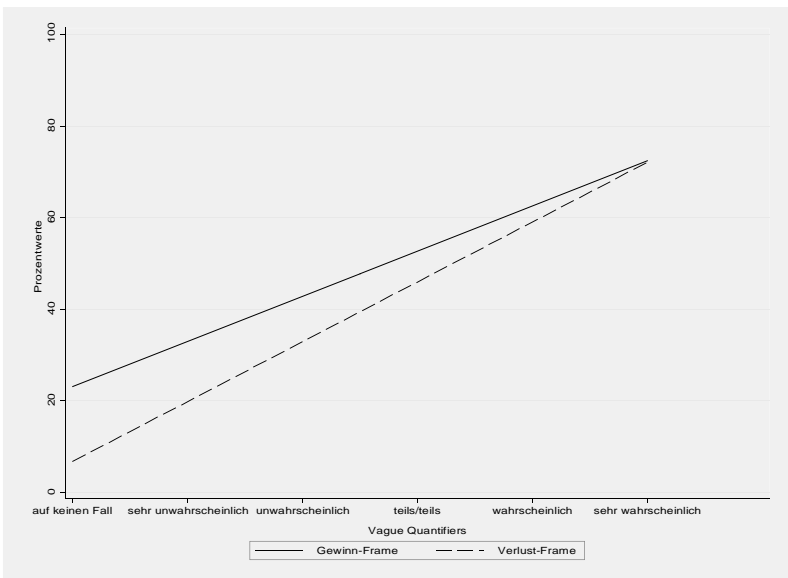


Abbildung 4 Inter-personelles Framing, Delikt: Verkehrsunfall (Frage 9)



Die in den Grafiken abgebildeten Response Functions der Befragten des Gewinn-Frames liegen bei allen Delikten über den Response Functions des Verlust-Frames. Bei den schwereren Delikten ‚Diebstahl‘ und ‚Raub‘ beträgt der Abstand zwischen den beiden Response Functions deutliche 20–25%. Somit liegt bei diesen beiden Delikten ein insgesamt recht starker Effekt im Sinne unserer Hypothese vor. Die Response Functions beim Delikt ‚Verkehrsunfall‘ liegen dagegen näher beieinander.¹¹

Ein noch stärkerer Effekt zeigt sich beim *intra-personellen Test*. Dabei schätzte dieselbe Person ihre subjektive Wahrscheinlichkeit, Opfer eines Raubs zu werden, jeweils zweimal ein: Durch Randomisierung wurden 2 Gruppen gebildet. Dabei beantworteten die Befragten in der ersten Gruppe zunächst die Frage nach der Wahrscheinlichkeit ‚ausgeraubt zu werden‘ (Frage 1, Verlust-Frame). Nach einigen weiteren Fragen wurden sie schließlich aufgefordert, die Frage nach ihrer subjektiven Wahrscheinlichkeit ‚nicht ausgeraubt zu werden‘ zu beantworten (Frage 9, Gewinn-Frame). Die Befragten der zweiten Gruppe schätzten dagegen die subjektive Wahrscheinlichkeit ‚ausgeraubt zu werden‘ zweimal ein (Frage 1 und 9 jeweils im Verlust-Frame). Die Ergebnisse dieser Analysen sind in Abbildung 5 und 6 visualisiert.

Im *intra-personellen Test* bewertete dieselbe Person die *Vague Quantifiers* im Gewinn-Frame höher als zuvor im Verlust-Frame. Das bestärkt die Ergebnisse des *inter-personellen Tests*. Dagegen gaben Respondenten in der Kontrollkondition, die zweimal dieselbe Frage beantworteten, relativ konsistente Antworten.

Gemäß unseren Erwartungen halten wir fest, dass bei häufigen Ereignissen den *Vague Quantifiers* höhere Prozentwerte zugeordnet werden. Umgekehrt werden bei seltenen Ereignissen geringere Werte auf der prozentualen Skala angegeben. Die Response Functions verschieben sich somit mit steigender Häufigkeit nach oben.

11 Eine detaillierte Darstellung der Kennwerte der OLS-Regressionen (Regressionskoeffizienten, Standardfehler, etc.) kann bei Bedarf von den Autoren angefordert werden.

Abbildung 5 Intra-personelles Framing, Delikt: Raub
(Frage 1 im Verlust-Frame, Frage 9 im Gewinn-Frame)

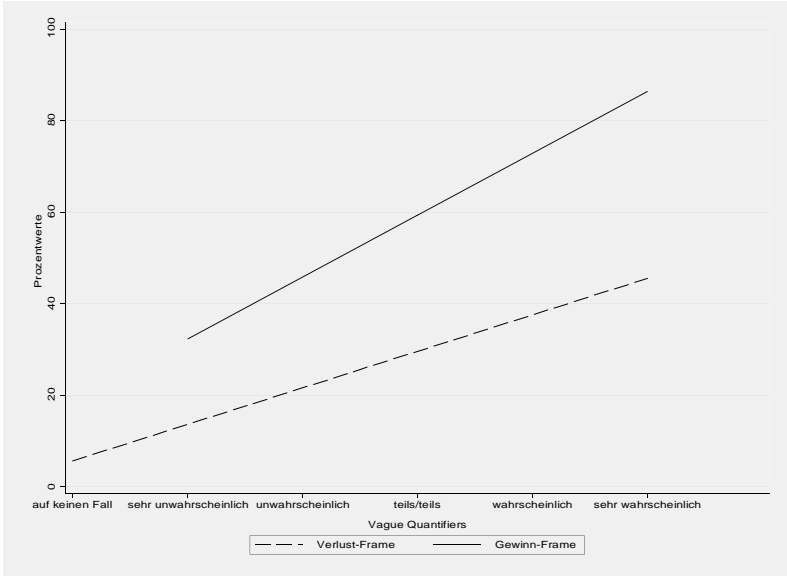
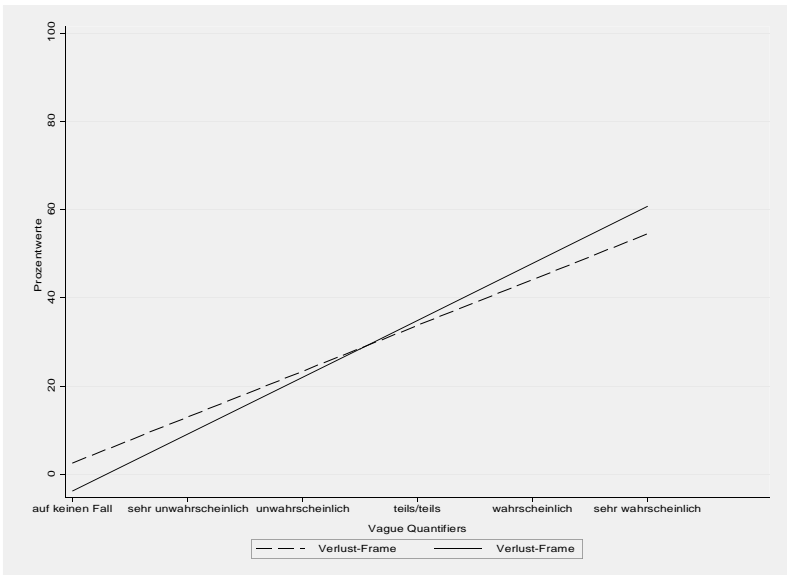


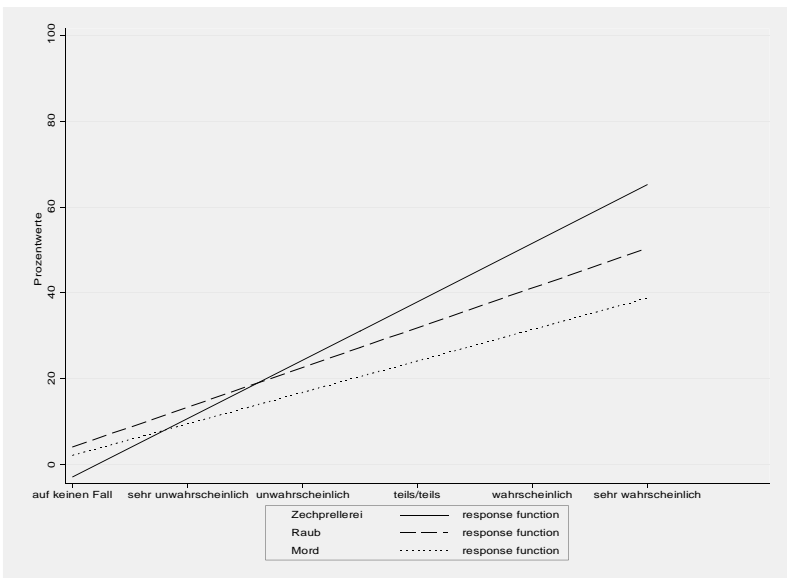
Abbildung 6 Intra-personelles Framing, Delikt: Raub
(Kontrollkondition: Frage 1 und 9 im Verlust-Frame)



4.3 Einfluss der Schwere einer Straftat

Hinsichtlich des Einflusses der Schwere einer Straftat auf das Mapping der verbalen mit der numerischen Wahrscheinlichkeitsskala bestätigt sich unsere Prognose: Mit zunehmender Deliktsschwere (aufsteigend: Zechprellerei, Raub, Mord) sinken die prozentualen Wahrscheinlichkeiten, die einem gegebenen *Vague Quantifier* zugeschrieben werden (vgl. Abbildung 7).

Abbildung 7 Deliktsspezifisches Mapping: Zechprellerei, Raub, Mord (Frage 1)



Aus Abbildung 7 wird ersichtlich, dass mit zunehmender Schwere eines Delikts die Steigungen der Response Functions abnehmen.¹² Vergleicht man die Wahrscheinlichkeiten zwischen verschiedenen Delikten, so wird deutlich, dass die verbalen Antwortkategorien nicht einheitlich verwendet werden. Unter ‚sehr wahrscheinlich‘ ermordet zu werden subsumieren Befragte eine kleinere Wahrscheinlichkeit,

12 Zechprellerei: $b = 13,6$; Raub: $b = 9,2$; Mord: $b = 7,3$. Weitere Kennwerte der Regressionen können bei Bedarf von den Autoren angefordert werden.

als unter ‚sehr wahrscheinlich‘ Opfer einer Zechprellerei zu werden. So liegt für das verbale Antwortlabel ‚sehr wahrscheinlich‘ beim Delikt ‚Mord‘ die mittlere prozentuale Wahrscheinlichkeit unter 40%. Dagegen liegt beim Delikt ‚Zechprellerei‘ für dasselbe verbale Antwortlabel die durchschnittliche zugeschriebene Wahrscheinlichkeit deutlich über 60%.

5 Diskussion

Bei der Interpretation von Daten, welche mit kategorialen Antwortskalen generiert wurden, sollte beachtet werden, dass der Fragekontext die Antworten beeinflusst. So können wir in einem randomisierten Experiment einen Effekt der Frageformulierung auf die subjektive Wahrscheinlichkeitsschätzung nachweisen: Respondenten schätzen ihre Viktimisierungswahrscheinlichkeit je nach Fragekontext bei formal gleichen Situationen unterschiedlich ein. Die subjektiven Wahrscheinlichkeiten unterscheiden sich, je nachdem, ob das Risiko eingeschätzt wird, von einer Straftat verschont zu bleiben oder einer Straftat zum Opfer zu fallen. Zudem zeigen wir, dass die gleichen verbalen Wahrscheinlichkeitsangaben in verschiedenen Situationen unterschiedliche Bedeutungen aufweisen. Einerseits ist mit ‚wahrscheinlich‘ bei seltenen Ereignissen ein geringeres Risiko gemeint als bei häufigeren Ereignissen. Darüber hinaus bedeutet eine ‚wahrscheinliche‘ Viktimisierung bei schweren Delikten ein geringeres Risiko als bei leichteren Delikten.

Ein direkter Vergleich von Anteilswerten über verschiedene Viktimisierungen hinweg ist somit als problematisch einzustufen, da die Interpretation der Antwortkategorien abhängig ist von der Frageformulierung, der Delikthäufigkeit und der Deliktsschwere. Wenn Befragte beispielsweise mit der Kategorie ‚sehr wahrscheinlich‘ ihr subjektives Risiko kommunizieren, Opfer einer bestimmten Straftat zu werden, dann meinen sie beim Delikt Mord eine geringere Wahrscheinlichkeit, als beim Delikt Raub. Um jedoch verbale Wahrscheinlichkeitsangaben zu verschiedenen Viktimisierungen direkt vergleichen zu können, müssten Befragte die verbalen Antwortkategorien für alle Delikte einheitlich verwenden: ‚Sehr wahrscheinlich‘ ausgeraubt zu werden sollte demnach genauso wahrscheinlich sein wie ‚sehr wahrscheinlich‘ ermordet zu werden. Da diese Annahme aber nicht zutrifft, lassen sich individuelle Unterschiede in der Risikoeinschätzung häufig nicht erkennen, wenn man lediglich Antworten auf den verbalen Antwortskalen vergleicht.

Da Befragte sich in ihren Antworten durch den Fragekontext beeinflussen lassen, besteht berechtigter Zweifel an der Annahme, dass die untersuchten Items

zur Risikoeinschätzung ein eindimensionales Konstrukt erfassen.¹³ Zu einer ähnlichen Schlussfolgerung kommt auch Kreuter (2002: 228):

„So nannten Befragte bei den Überlegungen zu ihren eigenen Risiken zunächst die allgemeinen Risiken, besannen sich auf ihre üblichen Alltagsroutinen und leiteten daraus ihre Antwort auf das eigene Risiko ab. Die explizite Erfassung allgemeiner Risiken vor der Abfrage subjektiver Risiken kann die Befragten möglicherweise um diese kognitive ‚Kalkulation‘ entlasten.“

Einen viel versprechenden Ausgangspunkt zur Erfassung dieser allgemeinen Risiken stellen frequentistische Ansätze zum Umgang mit Wahrscheinlichkeiten dar (vgl. Gigerenzer/Hoffrage 1995; Hoffrage et al. 2000).¹⁴

Neben diesen Überlegungen zu geeigneten Erfassungsstrategien von subjektiven Viktimisierungswahrscheinlichkeiten könnte als weiterer Anknüpfungspunkt an unsere Studie die Stärke der Framing-Effekte auf die verbale Skala und die numerische Skala verglichen werden. Einige Studien gehen davon aus, dass Kontext-Effekte auf der verbalen Skala stärker wirken als auf der prozentualen Skala (vgl. Moxey/Sanford 2000: 238; Welkenhuysen et al. 2001). Dagegen wird in anderen Studien berichtet, dass Framing-Effekte nur auf der numerischen Skala signifikant sind (vgl. Jasper et al. 2001). Eine vertiefte Untersuchung dieses Interaktionseffektes würde zu einem besseren Verständnis der Nutzung unterschiedlicher Antwortskalen beitragen.

- 13 Wir konnten in unseren Untersuchungen der Antwortstrategien von Befragten im Rahmen von qualitativen Interviews (N=12) im Pretest unseres Instruments feststellen, dass Viktimisierungswahrscheinlichkeiten bei häufigen Delikten (z. B. Zechprellerei) für das kommende Jahr aus Viktimisierungen des vergangenen Jahres abgeleitet wurden. Dagegen orientierten sich Befragte bei selteneren Delikten eher an der Viktimisierung anderer Personen. So wurden beispielsweise beim Delikt Mord zum Vergleich Zeitungsberichte herangezogen. Zukünftige Studien könnten diese Antwortstrategien noch differenzierter untersuchen.
- 14 „Frequentistische Ansätze“ basieren auf Erkenntnissen der Kognitionspsychologie, wonach die Darstellung und Kommunikation von statistischen Informationen in Form von Häufigkeiten intuitiver und verständlicher ist als alternative Darstellungsformen in Wahrscheinlichkeiten und Prozenten: "Natural frequencies facilitate inferences because they carry implicit information about base rates (...). They also correspond to the way in which humans have experienced statistical information over most of their history." (Hoffrage et al. 2000: 2261). Ein bekanntes Problem ist die subjektive Überschätzung von Risiken im Zusammenhang mit seltenen Ereignissen (fiktives Beispiel: „Wie wahrscheinlich ist es, dass eine zufällig ausgewählte Person aus ihrer Wohngegend innerhalb der nächsten 12 Monaten Opfer eines Raubs wird?“). Werden Items zur Messung subjektiver Risiken in Häufigkeiten formuliert, so fördert dies ein Denken in Verteilungen und führt in vielen Fällen zu realistischeren Schätzungen der Base Rate (fiktives Beispiel: „Stellen Sie sich 1000 Personen aus ihrer Wohngegend vor. Wie viele von diesen Personen werden innerhalb der nächsten 12 Monaten Opfer eines Raubs werden?“). Teigen (1974) zeigt, dass die subjektive Überschätzung von kleinen Wahrscheinlichkeiten abnimmt, wenn Befragte Risiken in Häufigkeiten anstatt Wahrscheinlichkeiten schätzen. Coutts (2002) zeigt, dass Häufigkeitsformate („distributional perspective“) zu realistischeren, niedrigeren Einschätzungen von deliktsspezifischen Viktimisierungsrisiken führen als andere Formate („singular perspective“).

Dieses bessere Verständnis würde dann eine begründete Auswahl einer verbalen oder numerischen Skala ermöglichen. So wird in der Medizin die Verwendung von verbalen Häufigkeitsausdrücken größtenteils abgelehnt (vgl. Nakao/Axelrod 1983). Entsprechend beschäftigen sich die meisten Fachartikel in diesem Bereich mit den Vor- und Nachteilen verschiedener Varianten von numerischen Skalen. Ghosh/Ghosh (2005) haben in einer umfassenden Review Ergebnisse von insgesamt 52 Studien zusammengestellt. Hierbei wurde in 17 Studien untersucht, ob Medizinstudenten und Ärzte numerische Wahrscheinlichkeitsformate richtig deuten und korrekt kommunizieren können:

„Physicians demonstrate widely varying understanding of probability terms and NNT [number needed to treat]“ (Ghosh/Ghosh 2005: 178).

Noch größere Probleme hinsichtlich des Verständnisses von numerischen Angaben wurden dabei in den restlichen 35 Studien auf Seiten der Patienten festgestellt.

Der direkte Vergleich von verbalen und numerischen Skalen führt allerdings in einigen Studien zu einer positiveren Bewertung der Ersteren. So schätzen Frauen ihr eigenes Brustkrebsrisiko auf einer verbalen Skala realistischer ein als auf einer numerischen Skala, bei der es zu einer deutlichen Überschätzung des objektiven Krebsrisikos kam (vgl. Woloshin et al. 1999).

Windschitl/Wells (1996: 359) kommen schließlich zu dem Ergebnis, dass sowohl numerische als auch verbale Maße ihre Berechtigung haben. Sie stellen fest:

„Results suggest that numeric measures tend to elicit deliberate and rule-based reasoning from respondents, whereas verbal measures allow for more associative and intuitive thinking“ (Windschitl/Wells 1996: 343).

Literatur

- ALLBUS, 1990 u. 2000: Codebücher (Mannheim: Zentrum für Umfragen, Methoden und Analysen, ZUMA). <http://www.gesis.org/Datenservice/ALLBUS/index.htm> (08.05.2007).
- Banks, S. M., P. Salovey, S. Greener, A. J. Rothman, A. Moyer, J. Beauvais und E. Epel, 1995: The effects of message framing on mammography utilization. *Health Psychology* 14: 178-184.
- Bradburn, N. und C. Miles, 1979: Vague Quantifiers. *Public Opinion Quarterly* 43: 92-101.
- British Crime Survey 2004 u. 2005: Fragebogen (London: Research development statistics, rds). <http://www.data-archive.ac.uk/doc/5347/mrdoc/pdf/5347userguide2.pdf> (08.05.2007).
- Budescu, D. und T. S. Wallsten, 1995: Processing linguistic probabilities. General principles and empirical evidence. S. 275-318 in: J. Busenmeyer, D. L. Medin und R. Hastie (Hg.): *Decision making from a cognitive perspective*. San Diego, CA: Academic Press.
- Conrad F. G., N. Brown und E. Cashman, 1998: Strategies for Estimating Behavioural Frequency in Survey Interviews. *Memory* 6: 339-366.

- Coutts, E., 2002: Context effects in the measurement of subjective probabilities in surveys. Universität Konstanz: Diplomarbeit Fachbereich Psychologie.
- Edwards, A., G. Elwyn, J. Covey, E. Matthews und R. Pill, 2001: Presenting Risk Information – A Review of the Effects of "Framing" and other Manipulations on Patient Outcomes. *Journal of Health Communication* 6: 61-82.
- Ghosh, A. K. und K. Ghosh, 2005: Translating evidence-based information into effective risk communication. Current challenges and opportunities. *Journal of Laboratory and Clinical Medicine* 145: 171-180.
- Gigerenzer, G. und U. Hoffrage, 1995: How to improve Bayesian reasoning without instruction. Frequency formats. *Psychological Review* 102: 684-704.
- Goocher, B. E., 1965: Effects of attitude and experience on the selection of frequency adverbs. *Journal of Verbal Learning and Verbal Behaviour* 4: 193-195.
- Grice, H. P., 1993: Logik und Konversation. S. 243-265 in: G. Meggle (Hg.): *Handlung, Kommunikation, Bedeutung*. Frankfurt am Main: Suhrkamp-Verlag.
- Gurm, H. S. und D. G. Litaker, 2000: Framing procedural risks to patients. Is 99% safe the same as a risk of 1 in 100? *Academic Medicine* 75(8): 840-842.
- Hakel, M. D., 1969: How often is often? *American Psychologist* 25: 533-534.
- Hammerton, M., 1976: How much is a large part? *Applied Ergonomics* 7: 10-12.
- Hoffrage, U., S. Lindsey, R. Hertwig und G. Gigerenzer, 2000: Communicating Statistical Information. *Science* 290: 2261-2262.
- Hörmann, H. 1983: The calculating listener, or how many are „einige, mehrere und ein paar“. S. 221-234 in: R. Bauerle, C. Schwarze und A. von Stechow (Hg.): *Meaning, use and interpretation of language*. Berlin: De Gruyter.
- Jasper, J., R. Goel, A. Einarnson, M. Gallo und G. Koren, 2001: Effects of framing on teratogenic risk perception in pregnant women. *The Lancet* 358: 1237-1238.
- Kahneman, D. und K. A. Tversky, 1979: Prospect theory: An analysis of decision under risk. *Econometrica* 47: 263-291.
- Karelitz, T. M. und D. Budescu, 2004: You Say "Probable" and I Say "Likely". Improving Interpersonal Communication with Verbal Probability Phrases. *Journal of Experimental Psychology Applied* 10: 25-41.
- King, G., J. L. Murray Christopher, J. A. Salomon und A. Tandon, 2004: Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research. *American Political Science Review* 98: 191-207.
- Kreuter, F., 2000: Uncertainty in capturing uncertainty. Toward measurement of fear of crime. Part I – Measuring subjective probability. (Girona, Annual Conference of the International Society of Quality of Life Studies). www.uni-konstanz.de/FuF/Verwiss/Schnell/Kreuter/girona.pdf (08.05.2007).
- Kreuter, F., 2002: *Kriminalitätsfurcht. Messung und methodische Probleme*. Opladen: Leske & Budrich.
- Kuhn, K. M., 1997: Communicating uncertainty. Framing effects on responses to vague probabilities. *Organizational Behavior and Human Decision Processes* 71: 55-83.
- Marteau, T. M., 1989: Framing of information. Its influence upon decisions of doctors and patients. *British Journal of Social Psychology* 28: 89-94.
- McNeil, B. J., S. G. Pauker, H. C. Sox und K. A. Tversky, 1982: On the elicitation of preferences for alternative therapies. *New England Journal of Medicine* 306: 1259-1262.
- Menon, G., P. Raghubir und N. Schwarz, 1995: Behavioral frequency judgments. An accessibility-diagnosticity framework. *Journal of Consumer Research* 22: 212-228.
- Merz, J. F., M. J. Druzdzel und D. J. Mazur, 1991: Verbal expressions of probability in informed consent litigation. *Medical Decision Making* 11(4): 273-281.

- Moxey, L. M. und A. J. Sanford, 1993: Prior expectation and the interpretation of natural language quantifiers. *European Journal of Cognitive Psychology* 5: 73-91.
- Moxey, L. M. und A. J. Sanford, 2000: Communicating quantities: A Review of Psycholinguistic Evidence of How Expressions Determine Perspective. *Applied Cognitive Psychology* 14: 237-255.
- Murakami, H., 2000: *South of the border, west of the sun*. New York: Vintage.
- Nakao, M. A. und S. Axelrod, 1983: Numbers are better than words – Verbal specifications of frequency have no place in medicine. *American Journal of Medicine* 74(6): 1061-1065.
- Newstead, S. E. und K. R. Coventry, 2000: The role of context and functionality in the interpretation of quantifiers. *European Journal of Cognitive Psychology* 12: 243-259.
- Pepper, S. und L. S. Prytulak, 1974: Sometimes frequently means seldom. Context effects in the interpretation of quantitative expressions. *Journal of Research in Personality* 8: 95-101.
- Peters, E., K. McCaul, M. Stefanek und W. Nelson, 2006: A Heuristic Approach to Understanding Cancer Risk Perception. Contributions from Judgement and Decision-Making Research. *Annals of Behavioral Medicine* 31: 45-52.
- Reuband, K. H., 2002: Subjektive Wahrscheinlichkeiten und Antwortmuster: Der Einfluss von Personenbezug und Skalierungsart. *ZA-Information* 50: 46-58.
- Reyna, V., 1981: The language of possibility and probability. Effects of negation on meaning. *Memory & Cognition* 9: 642-650.
- Schaeffer, N.C. 1991: Hardly ever or constantly? Group comparisons using vague quantifiers. *Public Opinion Quarterly* 55: 395-423.
- Schnell, R. und F. Kreuter, 2000: Das DEFECT-Projekt. Sampling-Errors und Nonsampling-Errors in komplexen Bevölkerungsstichproben. *ZUMA-Nachrichten* 47: 89-101.
- Schwarz, N., H. J. Hippler, B. Deutsch und F. Strack, 1985: Response categories. Effects on behavioural reports and comparative judgments. *Public Opinion Quarterly* 49: 388-395.
- Schwarz, N. und B. Scheuring, 1988: Judgments of relationship satisfaction. Inter- and intraindividual comparison strategies as a function of questionnaire structure. *European Journal of Social Psychology* 18: 485-496.
- Schwarz, N. und B. Scheuring, 1992: Selbstberichtete Verhaltens- und Symptommhäufigkeiten. Was Befragte aus Antwortvorgaben des Fragebogens lernen. *Zeitschrift für Klinische Psychologie* 22: 197-208.
- Simpson, R. H. 1944: The specific meanings of certain terms indicating differing degrees of frequency. *Quarterly Journal of Speech* 30: 328-330.
- Sutherland, H. J., G. A. Lockwood, D. L. Trichtler, F. Sem, L. Brooks und J. E. Till, 1991: Communicating probabilistic information to cancer patients – is there "noise" on the line? *Social Science & Medicine* 32(6): 725-731.
- Teigen, K. H., 1974: Overestimation of subjective probabilities. *Scandinavian Journal of Psychology* 15: 56-62.
- Tversky, K. A. und D. Kahneman, 1981: The framing of decisions and the psychology of Choice. *Science* 211: 453-458.
- Wänke, M., 2002: Conversational Norms and the Interpretation of Vague Quantifiers. *Applied Cognitive Psychology* 16: 301-307.
- Wallsten, T. S., D. Budescu, R. Zwick und S. M. Kemp, 1993: Preferences and reasons for communicating probabilistic information in numerical or verbal terms. *Bulletin of the Psychonomic Society* 31: 135-138.
- Weber, E. U. und D. J. Hilton, 1990: Contextual effects in the interpretation of probability words – perceived base rate and severity of events. *Journal of Experimental Psychology: Human Perception and Performance* 16(4): 781-789.

- Welkenhuysen, M., G. Evers-Kiebooms und G. D'Ydewalle, 2001: The language of uncertainty in genetic risk communication. Framing and verbal versus numerical information. *Patient Education and Counseling* 43: 179-187.
- Windschitl, P. D. und G. L. Wells, 1996: Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology-Applied* 2(4): 343-364.
- Wohlfahrtssurvey, 1998: Fragebogen (Mannheim: Zentrum für Umfragen, Methoden und Analysen, ZUMA). http://www.gesis.org/Dauerbeobachtung/Sozialindikatoren/Publikationen/Datenreport/pdf_files/ws1998.pdf (08.05.2007).
- Woloshin, S., L. M. Schwartz, W. C. Black und H. G. Welch, 1999: Women's perceptions of breast cancer risk. How you ask matters. *Medical Decision Making* 19(3): 221-229.
- Wright, D., G. Gaskell und C. O'Muirheartaigh, 1994: How much is "Quite a bit"? Mapping between numerical values and vague quantifiers. *Applied Cognitive Psychology* 8: 479-496.
- Zimmer, A. C., 1983: Verbal versus numerical processing of subjective probabilities. S. 159-182 in: R. W. Scholz (Hg.): *Decision making under uncertainty*. Amsterdam: Elsevier.

Korrespondenzadressen:

Ivar Krumpal
Universität Leipzig
Institut für Soziologie
Beethovenstraße 15
D-04107 Leipzig
krumpal@sozio.uni-leipzig.de

Heiko Rauhut
ETH Zürich
Professur für Soziologie
Universitätsstraße 41
CH-8092 Zürich
rauhut@gess.ethz.ch