
Soziale Probleme

Zeitschrift für soziale Probleme und soziale Kontrolle

23. Jahrgang 2012, Heft 2

Warum es sich lohnt, Alltagstheorien zum Strafen ernst zu nehmen Zur Vermittlung zwischen autoritären Einstellungen und Strafverlangen <i>Kurt Mühler und Christian Schmidtke</i>	133
Bestrafen wir Erkan härter als Stefan? Befunde einer experimentellen Studie <i>Ulrike Häßler und Werner Greve</i>	167
Hafterleben von Frauen mit Kindern Eine qualitative Fallstudie <i>Mechthild Bereswill und Julia Hellwig</i>	182
Der Kampf um Anerkennung in Gewaltkarrieren von Mädchen <i>Claudia Equit</i>	216
Die Vernachlässigung langfristiger Folgen in der Entscheidung zu abweichendem Verhalten Entwicklung und Gütekriterien einer neuen Messung von Selbstkontrolle <i>Sonja Schulz und Harald Beier</i>	251



CENTAURUS
Verlag & Media KG

ISSN 0939-608X

Bestrafen wir Erkan härter als Stefan?

Befunde einer experimentellen Studie

von Ulrike Häßler und Werner Greve

Zusammenfassung

Diese experimentelle Studie zu Strafeinstellungen unter Studierenden der Erziehungswissenschaft zeigt, dass in einem ‚Between-Subject Design‘ Personen mit einem typisch ausländischen Namen härter bestraft werden als Personen mit einem typisch deutschen Namen. Diese Effekte zeigen sich besonders in Fällen von schwerer Kriminalität wie Vergewaltigung und Totschlag. Um sozial erwünschtes Antwortverhalten auszuschließen, wurde der Fragebogen in drei Versionen konstruiert. Der punitive Gehalt der Antworten wurde mit einem speziellen Härtescore abgebildet. So konnte die subjektive Härte der verhängten Strafen für jeden Fall identifiziert werden. Die Ergebnisse dieser Studie sind vor allem eine Aufforderung zu weiteren Studien mit experimentellen Designs. So können subjektive Einstellungen der Befragten besser abgebildet werden als in traditionellen Designs.

1. Einleitung

Das Thema der „Ausländerkriminalität“ hat es von der wiederkehrenden Konjunktur zum Dauerbrenner geschafft: Wenn es kurzzeitig von aktuellen Themen aus den Schlagzeilen verdrängt wird, dann nicht, weil es endgültig erledigt wäre, sondern nur, um mit der nächsten Meldung zu neuer Sichtbarkeit zurückzukehren. Sicherlich hat sich die Konkurrenz durch andere dramatische – oder dramatisierte – Themen erhöht (Finanzkrisen, Umweltschäden globalen Ausmaßes oder terroristische Bedrohungen), aber nach wie vor wecken Meldungen von kriminellen Handlungen nicht-deutscher Täter hohe mediale Aufmerksamkeit. Es gibt empirische Hinweise darauf, dass solche Meldungen, wenn sie rezipiert werden, auch einen Einfluss auf strafbezogene

Einstellungen haben (Simonson 2009; Windzio et al. 2007); das gilt gerade auch in Bezug auf nicht-deutsche Tätergruppen (Windzio/Kleimann 2009). Der Umgang mit Straftätern, die nicht aus Deutschland kommen, wird seit längerer Zeit diskutiert (Johnson et al. 1995) mit durchaus konkreten politischen und rechtlichen Folgen: Die Einführung der Abschiebungsanordnung aufgrund einer „tatsachengestützten Gefahrenprognose“ im Zusammenhang mit Überwachungsmaßnahmen ausgewiesener Ausländer (§ 58a AufenthG) ist ein konkretes Beispiel.

Tatsächlich deuten verschiedene Befunde darauf hin, dass Ausländer anders und härter bestraft werden als Deutsche. In der Aktenanalyse von Schott et al. (2004) zeigte sich etwa, dass Ausländer ohne Aufenthaltsgenehmigung seltener zu Freiheitsstrafen auf Bewährung verurteilt wurden und meist längere Freiheitsstrafen erhielten (Pfeiffer et al. 2005: 83). Zwar unterschieden sich die Gruppen, die vergleichend untersucht wurden, nicht nur in Bezug auf den Migrantenstatus, sondern auch u.a. in Bezug auf Vorstrafen und Schwere der Schäden, die durch die Tat entstanden, dennoch kommen die Autoren zu dem Schluss, dass dieser Unterschied keine hinreichende Erklärung für die härtere Sanktionierung der untersuchten Migranten ist (Pfeiffer et al. 2005: 84). Allerdings haben Aktenanalysen mehrere methodische Schwierigkeiten. Die konfundierten Variablen lassen sich lediglich statistisch kontrollieren, was wenigstens bei einigen auch wahre Varianz auspartialisieren könnte. So könnten beispielsweise in unterschiedlichen Vorstrafenbelastungen bereits unterschiedliche Sanktionstendenzen enthalten sein. Nicht zuletzt in dem Umstand, dass die zugrundeliegenden Daten reine Hellfelddaten sind, könnte eine systematische Selektivität der Stichprobe dann enthalten sein, wenn sich auch die Anzeigebereitschaft im Hinblick auf gleich- oder fremdethnische Personen systematisch unterscheidet; aktuelle Daten aus repräsentativen Befragungen (Baier et al. 2009) deuten darauf hin. Aber natürlich werden – abgesehen von den justiziell vorgesehenen Aspekten wie Tatschwere etc. – zahlreiche weitere Faktoren die Strafzumessung beeinflussen; darauf deutet seit langem der bekannte Befund unterschiedlicher „regionaler Justizkulturen“ (Pfeiffer et al. 2005: 82) hin (vgl. auch Gabriel/Greve 1996).

Neben der skizzierten Hellfeldstudie, die faktische justizielle Strafzumessung auf korrelativem Wege (überwiegend auf der Basis von Akteninformationen) untersucht, sind vor allem Studien zu Strafeinstellungen („Punitivität“) und deren Einflussfaktoren durchgeführt worden (Gabriel/Greve 2008;

als Überblick Oswald/Bienek/Hupfeld-Heinemann 2009). Durch den Einsatz von Fallvignetten können schließlich auch konkrete (fiktive) Strafreaktionen zum einen auf den Einfluss (generalisierter) strafbezogener Einstellungen zurückgeführt werden, zum anderen auf tat- und täterbezogene Aspekte (Gabriel/Greve 1996). Jedoch sind auch hierzu bislang weit überwiegend korrelative Studien vorgelegt worden; nur selten wurde das Potenzial experimenteller Designs genutzt, Kausalhypothesen systematisch zu prüfen. Eben dies ist das Anliegen der im Folgenden vorzustellenden Studie. Durch ein Design, in dem verschiedene Versionen des zu beurteilenden Falles verschiedenen Personen vorgelegt werden („between-subjects“-Design) können die gerade im Fall der Untersuchung schwer kontrollierbaren Tendenzen sozialer Erwünschtheit zur Erklärung von Unterschieden in den Reaktionen auf verschiedene Versionen ausgeschlossen werden.

2. Durchführung der Studie

In Anlehnung an methodische Diskussionen der Punitivitätsforschung (Simonson 2009; Suhling et al. 2005) können Strafreaktionen und -einstellungen vor allem auf zwei Wegen erfasst werden: Selbstauskunftsaussagen und Fallgeschichten. Dabei spielt es eine entscheidende Rolle, welche Ebene von subjektiver Strafhärte erfasst und erklärt werden soll. Psychologisch lassen sich im Hinblick auf Einstellungen, also auch Einstellungen zu (harten) Strafen, drei Ebenen unterscheiden. Zum einen wird die kognitive Komponente angesprochen, d.h. die inhaltliche Beurteilung einer Tat und denkbarer Sanktionen, zum zweiten die emotionale Ebene, die die durch die Tat oder die Sanktion(en) ausgelösten Gefühle betrifft, und zum dritten die behaviorale Ebene, die das konkrete Handeln (z.B. die tatsächliche Strafreaktion) beschreibt. Zwar hängen diese drei Ebenen nicht nur empirisch, sondern auch konzeptuell zusammen (Gabriel/Greve 1996; Greve 2002), jedoch zeigt sich, dass sie durchaus nicht immer konvergieren.

Für die Abfrage von Einstellungen auf der kognitiven Ebene haben sich Selbstauskunftsaussagen bewährt (z.B. „Es sollte weniger gestraft werden“). Sie sind nicht nur ökonomisch einsetzbar, sondern haben auch den Vorteil, leicht mit anderen Studien vergleichbar zu sein. Jedoch wird hier nach einer Einschätzung gefragt, die vom Kontext losgelöst verschiedene Assoziationen hervorrufen kann (Suhling et al. 2005: 205). Zwar kann korrelativ geprüft

werden, inwieweit derartige Einstellungen mit anderen Einstellungen (oder Eigenschaften) der Person zusammenhängen, die konkrete Strafreaktion jedoch kann so schwerlich untersucht werden. Die Erforschung von Einstellungen auf der affektiven Ebene stellen eine besondere Herausforderung dar, für die bisweilen keine zufriedenstellende Lösung gefunden wurde (vgl. Suhling et al. 2005: 205; weiterführend Stalans 2009).

Fiktive Fallvignetten eröffnen den Zugang zur behavioralen Ebene von Einstellungen – oder vielmehr zu ihren Verhaltensintentionen. Durch Vignetten werden Strafsentscheidungen auf eine Ebene transferiert, in der die Reaktionen der Befragten neben der kognitiven Bewertung von Sachverhalten auch als „vorgestelltes Verhalten betrachtet werden“ kann (Suhling 2005: 9). Suhling et al. (2005) identifizierten verschiedene Vorteile dieser Methode: Zum einen sind Vignetten, anders als globale Items, konkret und spezifisch. Somit werden Antwortverzerrungen durch unterschiedliche Assoziationen und Verzerrungen durch unterschiedliche Reihenfolgen der Fragen in einem Fragebogen minimiert (zum Reihenfolgeeffekt Kury et al. 2004: 58). Des Weiteren eignen sich fiktive Fallgeschichten für experimentelle Studien, eben weil systematisch Merkmale und Fälle variiert werden können. Werden hinreichend viele und hinreichend verschiedene Vignetten vorgegeben, können die individuellen Reaktionen auf sie zugleich zu individuellen Strafhärtescores verrechnet werden (Gabriel/Greve 1996), vor allem dann, wenn hierzu die individuelle „Kalibrierung“ der konkreten Strafreaktionen erfasst wurde (Stiftung Universität Hildesheim 2008; Suhling et al. 2005: 208).

Die hier zu beschreibende Studie greift diese Überlegungen auf. In zwei Schritten wurden die Teilnehmer und Teilnehmerinnen dieser Studie zunächst gebeten, sich in Bezug auf mehrere konkrete Delikte (Fall-Vignetten) jeweils aus einer umfangreichen (für alle Delikte gleichen) Auswahl von Sanktionen für eine zu entscheiden; dann sollen alle in dieser Auswahl vorgegebenen Sanktionen in Bezug auf ihre von der beurteilenden Person empfundene Härte beurteilt werden (Einschätzung zwischen 1 und 100). So können individuell unterschiedliche Kalibrierungen (verschiedene Personen können gleiche Sanktionen als verschieden „hart“ empfinden) in der Auswertung berücksichtigt werden. Tatsächlich haben Studien gezeigt (vgl. Tremblay 1988; Suhling et al. 2005), dass sich eine personenunabhängige hierarchische Rangordnung von Sanktionen nicht bilden lässt (es ist nicht zu klären, ob die Bewährungsstrafe von einem halben Jahr „objektiv“ schwerwiegender ist als

eine Geldbuße über 5.000 Euro). Tremblay fand heraus, dass die Beziehung zwischen dem objektiven Strafmaß und der subjektiven Härte der Strafe als nichtlineare Funktion beschrieben werden kann (Tremblay 1988: 236).

Ziel der vorliegenden Studie ist nicht die Untersuchung der Frage, ob faktisch unterschiedlich auf verschiedene Fallkonstellationen reagiert wird (das wäre u.U. für professionell Strafende, z.B. Richter, interessant), sondern ob in Abhängigkeit von sachlich irrelevanten Fallmerkmalen unterschiedlich hart reagiert wird. Dazu war die Berücksichtigung der je individuellen Maßstäbe unerlässlich.

2.1. Beschreibung des Designs

Im Mittelpunkt der Untersuchung steht die Frage, ob deutsche und nicht-deutsche Täter verschieden hart bestraft werden. Der ethnische Hintergrund des Täters wurde hierzu in der Vignette nicht explizit benannt, sondern über die Variation des Vornamens nahe gelegt. Ausgangspunkt hierfür ist der häufiger replizierte Befund, dass Vornamen unterschiedliche Beurteilungstendenzen aktivieren. Beispielsweise hat eine aktuelle Studie der Universität Oldenburg (Spiegel-Online 2010) gezeigt, dass die Konnotation verschiedener männlicher Vornamen (z.B. Kevin, Maurice oder Justin versus Maximilian, Alexander oder Jakob) einen Einfluss auf das Bewertungsverhalten von Rechtschreibfehlern haben. Das ist umso bemerkenswerter, als die individuellen Assoziationen mit bestimmten Vornamen durchaus nennenswert variieren können (in der Studie konnte der Effekt für weibliche Vornamen nicht gefunden werden).

Entscheidend für die hier präsentierte Studie ist, dass zwar jede Person verschiedene Fallvignetten zu beurteilen hatte, die sich hinsichtlich verschiedener Merkmale des Täters und der Tat unterschieden, der kritische Vergleich aber nicht zwischen den Vignetten eines Fragebogens („within-subject“), sondern zwischen Varianten der jeweiligen Vignette in verschiedenen Fragebogenversionen („between-subject“) erfolgte. Diese Varianten unterschieden sich jeweils nur in einem kritischen Merkmal (z.B. Vornamen des Täters). Die Fragebogenversionen wurden unter den Teilnehmern zufällig verteilt (Fragebogenvariante I wurde 120mal, Variante II 127mal und Variante III 112mal bearbeitet); den Teilnehmern und Teilnehmerinnen war nicht bekannt, dass es mehrere Versionen des Fragebogens gab.

Von jeder der drei Fallgeschichten¹ (auf jeweils einer A4- Seite präsentiert) wurden dementsprechend zwei Versionen erstellt. In der nachfolgenden Übersicht wird gezeigt, welche Fragebogenvariante welche Merkmale enthält. Diese jeweils aufgeführten Merkmale sind der einzige Unterschied der verschiedenen Fragebogenvarianten. Gegebenenfalls auftretende (Mittelwerts- und Variations-) Unterschiede zwischen den jeweiligen Bedingungen (d.h. Gruppen, die eine Version bearbeitet haben) in der Strafhärtereaktion auf den jeweiligen Fall müssen also ausschließlich auf das variierte Merkmal zurückzuführen sein.

Übersicht 1: Aufbau des Experimentes²

Delikt	Variante I	Variante II	Variante III
Autodiebstahl	Täter: Kai R.	Täter: Hassan V.	Täter: Kai R.
Vergewaltigung	Täter: Erkan M.	Täter: Stefan L.	Täter: Erkan M.
Totschlag	Täter: Martin K.	Täter: Vitali W.	Täter: Vitali W.

Nach der Beschreibung der kurzen Fallgeschichte wurden 19 Sanktionsmöglichkeiten von Verwarnungen über soziale Dienste, Täter-Opfer-Ausgleich, Erziehungscamp, Bewährungsstrafen zwischen einem halben und mehr als zwei Jahren und Gefängnisstrafen von unter einem Monat bis fünf Jahre bis lebenslänglich (mit anschließender Sicherungsverwahrung) vorgegeben. Es sollte im Falle einer möglichen Kombination von Sanktionen (bspw. Geldbuße und soziale Dienste) die jeweils Härteste angegeben werden. Erst nach Abschluss dieses kompletten Fragebogenteils wurden die Teilnehmer/innen auf der nächsten Seite dazu aufgefordert: „Bitte ordnen Sie nun den folgenden Strafvarianten die Zahlen von 1 bis 100 zu. Die 1 entspricht dabei einer ‚sehr schwach empfundenen Bestrafung‘, die 100 einer ‚sehr stark empfundenen Bestrafung‘. Sie kennen den Strafenkatalog bereits aus der Beantwortung der Fälle.“ Für jede der 19 Sanktionen sollte eine Zahl von 1 bis 100 angegeben werden, je nachdem, wie hart die Strafe empfunden wird. Das Ziel war es, die subjektiv empfundene Strafhärte der 19 Sanktionen zu ermitteln, um dann in Bezug auf die einzelnen Delikte zu prüfen, wie hart (unabhängig von der konkreten Sanktion) jeweils gestraft wurde.

Die Darstellung der Ergebnisse bezieht sich dementsprechend auf die zwei Ebenen: Zum einen werden die konkreten Sanktionen für diesen Fall dargestellt; wie eben angesprochen sind hier auftretende mögliche Differenzen jedoch wegen unterschiedlicher individueller Kalibrierung schwer interpretierbar. Zum zweiten wird die subjektive Härte der gewählten Sanktion mit Hilfe des Strafhärtescores (1-100) in Bezug zu den unterschiedlichen in den Vignetten verwandten Vornahmen gesetzt.

Um die verschiedenen Gruppen berücksichtigen zu können und die maximale Teststärke zu erhalten, wurden Varianzanalysen gerechnet (Rasch/Friese 2007: 5). Da es drei Fragebogenvarianten gab (in denen jedoch nur zwei Vignettenversionen vorgegeben waren (siehe Übersicht 1), lässt sich so gleichzeitig abschätzen, ob sich auch gleiche Varianten unterscheiden (dies wäre ein Hinweis auf mangelhafte Reliabilität dieser Fallvignette).³ Die Varianzanalyse ist vergleichsweise robust gegen die Verletzung von Voraussetzungen (Varianzhomogenität, Normalverteilung). Während Varianzhomogenität für die hier mitgeteilten Befunde i. d. R. vorliegt, ist die Normalverteilungsannahme für die einzelnen Vignetten fast immer verletzt, wie bei Urteilen zu unterschiedlichen Fallvorgaben anhand des gleichen Reaktionsmaßstabs nicht anders zu erwarten.

2.2. Stichprobe

Im Rahmen eines Seminars der Universität Hildesheim wurden (2007) Bekannte und Angehörige von Studierenden der Erziehungswissenschaft befragt. Insgesamt beteiligten sich 359 Personen, von denen etwas über die Hälfte (51,8 %) zwischen 21 und 30 Jahren alt waren. Im Mittel liegt das Alter der Befragten bei 32 Jahren (Standardabweichung: 13,7 Jahre). Das Verhältnis zwischen männlichen und weiblichen Befragten ist nahezu ausgeglichen: 54,5 Prozent der Befragten waren weiblich, 45,5 Prozent männlich. Allerdings zeigten sich im Gegensatz zu anderen Studien in Bezug auf die Strafhärteeinschätzungen keine signifikanten Unterschiede zwischen den Geschlechtern (Gabriel/Greve 1996).

Bei den Befragten handelte es sich fast ausschließlich um deutsche Befragte (94,3%), lediglich 3,6 Prozent gaben ein anderes Geburtsland an, das in den meisten Fällen im früheren Ostblock lag (Russland, Polen, Kasachstan). Das Bildungsniveau der Befragten war relativ hoch: über die Hälfte

(56,8%) gaben Abitur oder akademische Abschlüsse als höchsten Bildungsstand an. Es handelt sich damit um eine selektive Gelegenheitsstichprobe. Sie kann zwar keinen Anspruch auf Repräsentativität erheben, dürfte aber im Sinne der Fragestellung eine eher konservative Selektion darstellen: Aus Forschungen zur Einstellung zur Todesstrafe wurde ermittelt, dass mit einer niedrigeren Bildung ein erhöhtes Strafbedürfnis einhergeht (Bohm 1987). Da in dieser Stichprobe hauptsächlich Personen mit einem hohen Bildungsniveau befragt wurden, ist ein geringeres Strafbedürfnis zu erwarten. Urteilsrelevante Vorurteile (zumal in Bezug auf Strafhärte) sollten in der so erreichten Population somit eher unterdurchschnittlich auftreten – die Hypothese der Ungleichbehandlung verschiedener Ethnien hat es also vergleichsweise schwerer.

3. Ergebnisse

Fall 1 (Autodiebstahl)

Die erste Vignette, die genauer auf die unterschiedlichen Strafhärten aufgrund des Vornamens untersucht werden soll, ist Vignette 1: der Autodiebstahl. Die genaue Formulierung lautete:

Der 17-jährige Kai R. / Hassan V. wird wegen Diebstahl eines PKW festgenommen. Bei dem Auto handelt es sich um einen gebrauchten Mercedes C-Klasse (Wert 18.750 €).

Die Sanktionen, die in diesem Fall gewählt worden, waren zu 30 Prozent die sozialen Dienste gefolgt vom Täter-Opfer Ausgleich (16%), Erziehungscamp (10%) und Bewährungsstrafe von einem halben Jahr (10%). Die konkrete Strafauswahl der eher resozialisierenden Sanktionen unterscheidet sich kaum in den zwei Varianten: in dem Fall Hassan wurde etwas öfter Bewährungsstrafe (13,5%) von einem halben Jahr und seltener der Täter-Opfer Ausgleich (9,5%) gewählt.

Der Vergleich der drei Fragebogenvarianten in Bezug auf die Strafhärte ergab eine Tendenz zu mehr subjektiver Strafhärte bei dem ausländischen Täter (Hassan V.; Variante 2). Dieser Unterschied ist jedoch in der Grundgesamtheit nicht signifikant ($\alpha=0,489$, $F=0,826$, $df=2$). Es kann somit nicht ausgeschlossen werden, dass der deskriptiv erkennbare Unterschied zwischen

dem Mittelwert der „Kai“-Versionen I und III (31,1) und dem der „Hassan“-Version II (33,9) zufällig entstand.

Fall 2 (Vergewaltigung)

Ein deutlich gravierenderes Delikt wird in Vignette 2 beschrieben. Der genaue Wortlaut dieser Fallgeschichte lautete:

Auf dem spätabendlichen Nachhauseweg wird die Schülerin Maria S. (18) von dem 20-jährigen Erkan M. / Stefan L. mit freundlichen Komplimenten überhäuft. Nachdem sie ihn mit dem Hinweis, dass sie bereits einen Freund habe, abweist, wird er handgreiflich. Es kommt zur Vergewaltigung.

74,3 Prozent der Befragten würde hier eine Gefängnisstrafe von einem halben bis fünf Jahre verhängen. Über 20 Prozent würden eine lebenslängliche Haftstrafe oder eine Kombination dieser mit Sicherungsverwahrung verhängen. Über 71 Prozent der Befragten würde Stefan L. eine Gefängnisstrafe zwischen einem halben und fünf Jahren geben. Aber nur 60 Prozent gaben eine solche Strafe für Erkan M. an. Der Unterschied zeigt sich ebenfalls in der Anwendung der zwei Strafen „Lebenslänglich“ und „Lebenslänglich mit Sicherungsverwahrung“: 14,2 Prozent wählte diese Sanktionen bei Stefan L. und 24,6 Prozent bei Erkan M. Das deutet schon bei der Auswahl der Sanktionen an, dass repressivere Sanktionen im Falle eines ausländischen Vornamens angewandt werden könnten.

Diese Tendenz bestätigt sich in der Varianzanalyse des Strafhärtescore. Es zeigt sich, dass im Falle des deutschen Vornamens (Stefan L.) weniger subjektive Strafhärte der einzelnen Sanktionen angegeben wurde. In Variante I wurde eine durchschnittliche Strafhärte von 69,6 angegeben, für Variante II im Mittel 61,7 und für Variante III 71,8. Zieht man nun die beiden Versionen zusammen, die einen ausländischen Vornamen in der Fallbeschreibung hatten (70,7), ergibt sich ein Unterschied zur deutschen Variante von 9 Punkten auf der Strafhärteskala (1-100). Diese Differenz ist zudem statistisch signifikant ($\alpha=0,002$, $F=6,434$, $df=2$). Nach Rasch/Friese (2007) bedeutet ein signifikantes Ergebnis in der Varianzanalyse lediglich, dass sich die Gruppen in ihren Mittelwertausprägungen unterscheiden. Um die genauen Unterschiede zwischen den Gruppen zu ermitteln, dienen Post-hoc Analysen. Diese bestätigen

in diesem Fall, dass eine signifikante Abgrenzung zwischen den Varianten „Erkan“ und „Stefan“ vorliegt.

Fall 3: Totschlag

In Vignette 3 ist eine weitere gravierende Straftat beschrieben. Der genaue Wortlaut im Fragebogen ist:

Es kommt zu einem heftigen Streit zwischen Martin K. / Vitali W. (23) und seiner Freundin Christina B. Dem vorausgegangen waren schwere Untreuevorwürfe gegen die 20-jährige. Bei der Auseinandersetzung tötet Martin / Vitali seine Freundin im Affekt. Als Grund gibt er später vor dem Landgericht Hildesheim Eifersucht an.

Für die Tötung im Affekt wurde zu 52 Prozent die Strafe lebenslang gewählt (Haftstrafe von 5 Jahren: 26,8%, Gefängnisstrafen von weniger als 5 Jahren: 5,1%). Für Vitali wurde häufiger die Strafe „Lebenslänglich mit Sicherungsverwahrung“ gewählt als für Martin (15,5% zu 10,9%).

Dementsprechend zeigt sich für Variante eins eine durchschnittliche Strafhärte von 81,1, für Variante II eine Strafhärte von 81,8 und für die Variante mit dem ausländischem Vornamen eine Strafhärte von 87,4. Der Unterschied zwischen der Version mit deutschem und „ausländischem“ Vornamen beträgt also 5,9 Punkte auf der Strafhärteskala. Wegen fehlender Homogenität der Varianzen wurde hier ein Welch-Test gerechnet, der den gefundenen Unterschied als signifikant ausweist ($\alpha < 0,01$).

Der Befund spricht dafür, dass Täter mit ausländischem Vornamen mit als härter empfundenen Strafen sanktioniert werden, sofern es sich um ein gravierendes Delikt handelt. Eine Interpretation des Befundes, dass hier eine Vergewaltigung deutlichere Unterschiede produzierte als eine Tötung ist mit den vorliegenden Daten nicht valide möglich, da unterschiedliche Testverfahren verwandt wurden (Varianzanalyse bzw. Welch-Test). Deskriptiv zeigt sich für den Fall der Vergewaltigung ein Unterschied zwischen den Vornamen von 9 Punkten auf der Strafhärteskala; für den Fall des Totschlages ergibt sich ein Unterschied von 6 Punkten.

4. Diskussion

Der zentrale Befund der vorliegenden Studie ist es, dass in einem „between-subject“ Design identische Fallkonstellationen mit gravierenden Delikten (Verbrechen) dann systematisch verschieden beurteilt (sanktioniert) werden, wenn der Vorname des Täters auf einen fremdethnischen Hintergrund hindeutet. Wichtig ist dabei der Befund, dass nicht nur die faktische („objektive“) Sanktion, sondern insbesondere auch die je individuell empfundene Härte der Reaktion zwischen den Versionen variiert – stets zu Ungunsten des nicht-deutschen Täters. Die Selektivität der Stichprobe schränkt diesen Befund aus zwei Gründen nur unwesentlich ein. Zum einen dürfte die Selektion, wie ausgeführt, im Sinne der Fragestellung eher konservativ gewirkt haben (zuungunsten eines Unterschiedes), zum anderen wäre der Befund sozial auch dann relevant, wenn er nur in Teilen der Gesamtpopulation gefunden würde.

Die experimentelle Studie weist darauf hin, dass Täter mit einem „ausländischen“ Vornamen härter bestraft werden. Wir verhängen andere Sanktionen, und zwar solche, die wir als subjektiv härter empfinden. Als mögliche Erklärungen könnten verschiedene Theorien herangezogen werden. Es ist – empirisch wie theoretisch – nicht klar, ob dieser Befund auf unterschiedliche Einstellungen gegenüber Angehörigen anderer Ethnien schließen lässt. Empirisch sind Einstellungen in der vorliegenden Studie nicht unabhängig von den Sanktionsreaktionen erfasst worden; theoretisch ist unklar, ob das vorgestellte Verhalten *Komponenten* oder *Konsequenzen* von Einstellungen sind (Stalans 2009; vgl. auch Greve 2002); so könnte man den Befund als Beleg für unterschiedliche Einstellungen lesen – oder ihn (bei Vorliegen weiterer Daten) damit erklären.

Für Unterschiede in der Behandlung von Tätern, die ethnischen Minderheiten zugeordnet werden können und deutschen Delinquenten könnten auch der Ansatz der sozialen Identität (Hoog/Abrams 1988) und das Modell der subjektiven Gruppendynamik (Abrams 2000) relevant sein. Danach existieren spezielle Mechanismen, die den Unterschied zwischen der eigenen und der fremden Gruppe steigern. Die Zugehörigkeit zur eigenen Gruppe bildet unter anderem die Grundlage für die Identitätsentwicklung und die Weitergabe gesellschaftlicher Normen und Werte (van Prooijen 2009: 213). Auch hier

wären zusätzliche Daten erforderlich; dabei ließe sich dieser Faktor auch experimentell variieren.

Weitere mögliche Erklärungen sind, dass Strafreaktionen generell eher auf einer affektiven als auf einer behavioralen Ebene zu verorten sind (van Prooijen 2009: 223). Das heißt, selbst wenn z. B. belegt werden würde, dass es in der Quantität und Qualität von Straftaten keine Unterschiede zwischen Tätern aus ethnischen Minderheiten und deutschen Tätern gäbe, Vorurteile weiterhin bestehen bleiben könnten. Einstellungen sind womöglich nur schwer durch wissenschaftliche Belege zu verändern.

Als mögliche Erklärung dafür, dass Menschen aus anderen Gruppen härter bestraft werden, führt van Prooijen (2009) die Angst vor den Mitgliedern aus fremden Gruppen an und resümiert, dass solche Befragungen vor allem die intuitiven und emotionalen Reaktionen auf Verbrechen erfasst. Weiterführend kann die Ausgrenzung bestimmter Gruppen oder Menschen aus geltenden Normen und Werten als „moralische Exklusion“ bezeichnet werden (Optow 1990: 1). Es wird davon ausgegangen, dass die Reaktionen auf ein delinquentes Verhalten zunächst immer intuitiv/emotional sind und diese Entscheidung später mit rationalen Argumenten begründet wird (Löbmann et al. 2007: 10). Zum selben Ergebnis kamen auch internationale Studien zur Todesstrafe: die Befürwortung dieser Sanktion ist im höchsten Maße durch Emotionen geprägt. Diese „verzerrte Anpassung“ wird vor allem dadurch deutlich, dass selbst bei wissenschaftlichen Belegen darüber, dass es keine Abschreckungswirkung durch diese Sanktion gibt, dennoch der Anteil der Befürworter gleich bleibt (Bohm 1987: 389; Ellsworth/Ross 1983: 149).

Man kann argumentieren, dass die hier gefundene Differenz sozialer Reaktionen ihrerseits eine soziale (auch eine politische) Reaktion erfordert: *„differential treatment of offenders based on social categories can be considered as an instance of discrimination, which is a legal offense in itself in most democratic countries.“* (van Prooijen 2009: 226). Jedoch geht eine normative (z. B. sozialpolitische) Forderung über das hinaus, was empirische Befunde belegen und legitimeren können. Freilich: der Befund belegt die Notwendigkeit dieser Diskussion.

Anmerkungen

- 1 Die Studie untersuchte mehrere Hypothesen. Aus diesem Grund wurden verschiedene Merkmale variiert, die nach bisherigen Studien einen Einfluss auf die Strafzumessung haben: neben der ethnischen Zugehörigkeit auch das Motiv und die Vorstrafenbelastung. Die vorliegende Darstellung diskutiert nur den ersten Aspekt (Strafhärte in Abhängigkeit von der – nahegelegten – ethnischen Zugehörigkeit des Täters) – daher werden nur Ergebnisse für die drei diesbezüglich variierten Vignetten berichtet.
- 2 Weiterführend zu dieser Untersuchung Stiftung Universität Hildesheim, Institut für Psychologie 2008.
- 3 Beim paarweisen Vergleich über mehrere T-Tests geht die Teststärke wegen der alpha-Fehlerkumulierung verloren. Die Reliabilität wird dadurch deutlich, dass sich die Versionen mit den gleichen Geschichten nicht bedeutsam voneinander unterscheiden. Gerade die drei Versionen des Fragebogens machen den experimentellen Charakter dieser Studie deutlich.

Literatur

- Abrams, Dominic/Marques, José M./Bown, Nicola/Henson, Michelle, 2000: Pro-Norm and Anti-Norm Deviance Within and Between Groups. *Journal of Personality and Social Psychology* 78: 906-912.
- Baier, Dirk/Pfeiffer, Christian/Simonson, Julia/Rabold, Susanne, 2009: Jugendliche in Deutschland als Opfer und Täter von Gewalt. Hannover: KFN-Forschungsbericht.
- Bohm, Robert M., 1987: American Death Penalty Attitudes. A Critical Examination of Recent Evidence. *Criminal Justice and Behavior* 14: 380-396.
- Ellsworth, Phoebe C./Ross, Lee, 1983: Public Opinion on an Capital Punishment. A Close Examination of the View of Abolitionists and Retentionists. *Crime & Delinquency* 29: 116-169.
- Gabriel, Ute/Greve, Werner, 1996. „Strafe muß sein!“ Sanktionsbedürfnisse und strafbezogene Einstellungen: Versuch einer systematischen Annäherung. S. 185-214 in Pfeiffer, C./Greve, W. (Hrsg.), *Forschungsthema Kriminalität*. Baden-Baden: Nomos.
- Gabriel, Ute/Greve, Werner, 2008. Strafbedürfnisse und Strafeinstellungen. S. 467-476 in: Steller, M./Volbert, R. (Hrsg.), *Handbuch der Rechtspsychologie (Handbuch der Psychologie, Band 9)*. Göttingen: Hogrefe.
- Greve, Werner, 2002. Handlungstheorien. S. 300-325 in: Frey, D./Irlle, M. (Hrsg.), *Theorien der Sozialpsychologie, Band II (2. Auflage)*. Bern: Huber.
- Hoog, Michael A./Abrams, Dominic, 1988: *Social Identification: A Social Psychology of Inter-group Relations and Group Processes*. London: Routledge
- Johnson, James D./Whitstone, Erik/Anderson, Lee A./Gatto, Leslie, 1995: Justice is still not Colorblind. *Differential Racial Effects of Exposure to Inadmissible Evidence*. *Personality and Social Psychology Bulletin* 21: 893-898.
- Kury, Helmut/Kania, Harald/Obergfell-Fuchs, Joachim, 2004: Wortüber sprechen wir, wenn wir über Punitivität sprechen? Versuch einer konzeptionellen und empirischen Begriffsbestimmung. S. 51-88 in: Lautmann, R./Klimke, D./Sack, F. (Hrsg.), *Punitivität (Kriminologisches Journal, 8. Beiheft)*. Weinheim: Juventa.

- Löbmann, Rebecca/Suhling, Stefan/Greve, Werner, 2007: Emotionale Reaktionen auf Straftaten. Eine explorative Studie zu Unterschieden zwischen intuitiven und rationalen Strafurteilen. *Kriminologie und Rechtssoziologie* 1: 9-19.
- Opatow, Susan, 1990: Moral Exclusion and Injustice: An Introduction. *Journal of Social Issues* 1: 1-20.
- Oswald, Magrit E./Bieneck, Steffen/Hupfeld-Heinemann, Jörg (Hrsg.), 2009: *Social Psychology of Punishment of Crime*. Chichester, UK: Wiley-Blackwell.
- Pfeiffer, Christian/Kleinmann, Matthias/Petersen, Sven/Schott, Tilmann, 2005: Migration und Kriminalität. Ein Gutachten für den Zuwanderungsrat der Bundesregierung (Interdisziplinäre Beiträge zur kriminologischen Forschung Band 27). Baden-Baden: Nomos.
- Rasch, B./Friese, M./Hofmann, W./Naumann, E., 2006: *Quantitative Methoden* (2. Auflage) Heidelberg: Springer.
- Simonson, Julia, 2009: Punitivität: Methodische und konzeptionelle Überlegungen zu einem viel verwendeten Begriff. *Zeitschrift für Jugendkriminalitätsrecht und Jugendhilfe* 1: 112-119.
- Spiegel-Online 2010: Grundschullehrer-Vorurteile. Kevins bekommen schlechtere Noten (von O. Trenkamp). *Internetquelle*: [<http://www.spiegel.de/schulspiegel/wissen/0,1518,712948,00.html>].
- Stalans, Loretta J., 2009: Measuring Attitudes to Sentencing and Sentencing Goals. S. 231-254 in: Oswald, M.E./Bieneck, S./Hupfeld-Heinemann, J. (Hrsg.), *Social Psychology of Punishment of Crime*. Chichester, UK: Wiley-Blackwell.
- Stiftung Universität Hildesheim, Institut für Psychologie (Hrsg.), 2008: Warum strafen wir? Eine empirische Untersuchung zu Einflussfaktoren auf strafbezogene Einstellungen (Reihe Wissenschaft im Studium, Band 10). Hildesheim: Universitätsverlag.
- Suhling, Stefan/Löbmann, Rebecca/Greve, Werner, 2005: Zur Messung von Strafeinstellungen: Argumente für den Einsatz von fiktiven Fallgeschichten. *Zeitschrift für Sozialpsychologie* 36/4: 203-213.
- Tremblay, Pierre, 1988: On Penal Metrics. *Journal of Quantitative Criminology* 4/3: 225-245.
- Van Prooijen, Jan-Willem., 2009: Offenders' Social Categorization. Ingroup Bias or Black Sheep Effect? S. 211-230 in: Oswald, M.E./Bieneck, S./Hupfeld-Heinemann, J. (Hrsg.), *Social Psychology of Punishment of Crime*. Chichester, UK: Wiley-Blackwell.
- Windzio, Michael/Kleinmann, Matthias, 2009: Criminal Society and Criminal Immigrants. A Social Construction of Reality by Mass Media? S. 93-112 in: Oswald, M.E./Bieneck, S./Hupfeld-Heinemann, J. (Hrsg.): *Social Psychology of Punishment of Crime*. Chichester: Wiley-Blackwell.
- Windzio, Michael/Simonson, Julia/Pfeiffer, Christian/Kleinmann, Matthias, 2007: Kriminalitätswahrnehmung und Punitivität in der Bevölkerung. Welche Rolle spielen die Massenmedien? Ergebnisse der Befragung zu Kriminalitätswahrnehmung und Strafeinstellungen 2004 und 2006 (KFN-Forschungsberichte, Nr. 103). Hannover: KFN. *Internetquelle*: [<http://kfn.de/versions/kfn/assets/fb103.pdf>].

Harder Punishment to Foreigners? Results of an Experimental Study

Abstract

This experimental study on the penal attitudes of students of educational science shows that in a “between-subject” design persons with a typical non-German name are punished harder than persons with a typical German name. The results show this effect especially in cases of serious crimes like rape and homicide. To eliminate socially desirable response behaviour the questionnaire was constructed in three versions with different names. The punitive content of the answers was measured by a special score which allowed to identify the subjective severity of the punishment imposed for each case. The results are a demand for further studies with an experimental design. In this way the subjective attitudes of the respondents can be measured better than with a traditional design.

Ulrike Häßler

*Kriminologischer Dienst Celle
Fuhsestraße 30
29221 Celle*

ulrike.haessler2@justiz.niedersachsen.de

Werner Greve

*Universität Hildesheim
Institut für Psychologie
Marienburger Platz 22
31141 Hildesheim*

wgreve@rz.uni-hildesheim.de