

Running head: GEFÜHLE UND DEREN EINFLUSS AUF STRAFURTEILE

Der Einfluss von Gefühlen auf die Beurteilung von Straftaten

Corinne Gröbli, Sibylle Hafner, Marianne Spycher,

Bruno Sternath und Antonia Wolleb

Institut für Psychologie, Universität Bern, Schweiz

Korrespondenzadresse:

Bruno Sternath
Hombergstrasse 65, 4600 Olten
Schweiz
E-Mail: bruno.sternath@bluewin.ch
Tel. ++41 62 212 59 91
Fax ++ 41 62 212 59 91

Abstract

Ziel unserer Untersuchung war es, herauszufinden, ob Versuchspersonen die bei der Beurteilung normabweichenden Verhaltens durch eine gleichzeitig auszuführende Doppelaufgabe daran gehindert werden, ihre Gefühle zu kontrollieren, härtere Strafen fordern als Versuchspersonen, die nur mit der Beurteilung der Straftaten beschäftigt sind. 64

Versuchspersonen wurden vier Bedingungen zugeteilt. Die Unterschiede bestanden im Affektgehalt der Geschichten und einer manuellen Ablenkung (Tapping). Es gibt keinen signifikanten Unterschied mit und ohne Tapping in der Beurteilung des Strafmasses. In beiden Fällen wurde das Strafmass als leicht zu hoch eingeschätzt. Bei der Urteilsbedingung mit hohem Affektgehalt wird das Urteil als angemessener empfunden. Es besteht auch keine Interaktion. Die durchgeführten postexperimentellen Tests gaben keinen Aufschluss über relevante Störvariablen. Für differenziertere Resultate sollte man das Strafmass für die Tat der Vignetten neu definieren.

Ähnliche Straftaten werden vor Gericht und in der Bevölkerung als unterschiedlich schwer beurteilt und werden folglich unterschiedlich schwer bestraft. Worauf sind diese Unterschiede in der Beurteilung von Delikten zurück zu führen? Kriminologische Forschungen bestätigen, dass auch Laien bei der Beurteilung delinquenten Verhaltens rechtliche Kriterien berücksichtigen und ihre Urteile sorgfältig abgewägt fällen (durch Argumentation und Reflexion). Andere Studien (z.B. Applegate, Cullen, Link, Richards & Lanza-Kaduce 1996; Jacoby & Cullen, 1999; McKelvie, Mitchell, Arnott & Sullivan 1993; Miller, Rossi & Simpson 1986) konzentrieren sich auf die Stärke des Einflusses von sozioökonomischen Faktoren in der Bevölkerung (Alter, Geschlecht, Rasse, Bildung, Einkommen etc.) auf die Strenge der Beurteilung. Dabei stellte sich heraus, dass sich nur ein Teil der unterschiedlichen Beurteilung von Straftaten durch sozioökonomische Merkmale erklären liess. Aus dem Forschungsbericht von Haidt (2001) geht hervor, dass moralische Urteile vor allem aus dem emotionalen Erleben heraus gefällt werden. Löbmann, Suhling und Greve (2003) zeigten, dass der Schweregrad der Straftat, das Ausmass der Schädigung des Opfers sowie die Art der Beurteilung (emotional versus rational) einen signifikanten Einfluss auf die Beurteilung des Strafmasses haben. Im Zwei-Prozess Modell der Informationsverarbeitung (Eagly & Chaiken, 1993; Petty & Cacioppo, 1986) lassen sich die kognitiven und emotionalen Anteile des Strafurteils miteinander verknüpfen,

so dass Personen sich dann von ihren Gefühlen leiten lassen sollten, wenn sie nicht motiviert oder nicht in der Lage sind, Urteile rational zu fällen.

Unsere Forschungsfrage lässt sich wie folgt formulieren: Lassen sich Personen insbesondere dann von ihren Gefühlen leiten, wenn sie nicht in der Lage sind, Urteile rational zu fällen? Wir setzen voraus, dass das normabweichende Verhalten eine emotionale Reaktion auslöst.

Unsere Hypothese lässt sich wie folgt formulieren: Versuchspersonen (Vpn), die bei der Beurteilung normabweichenden Verhaltens durch eine gleichzeitig auszuführende Doppelaufgabe daran gehindert werden, ihre Gefühle zu kontrollieren, fordern härtere Strafen als Vpn, die nur mit der Beurteilung der Straftaten beschäftigt sind.

Methoden

Versuchspersonen

Bei den Versuchspersonen handelt es sich um Studentinnen und Studenten der Universität Bern. Die Stichprobe besteht aus 64 Vpn, davon 36 Frauen und 12 Männer unter 30, w 16 Frauen über 16 Jahre alt. Die Vpn sind mit der gerichtlichen Urteilsbildung nicht speziell vertraut, es sind Bekannte oder gehören zu unserer Studienrichtung. Die Teilnahme wird nicht entlohnt. Durch Quasirandomisierung werden die unter und über 30 Jährigen separat den Bedingungen zugeteilt, durch Zufallsauswahl entweder in eine Experimental- oder

Kontrollgruppe. Es resultieren vier Bedingungen, die Resultate von je 16 Vpn liefern.

Apparate und Materialien

Der gesamte Versuch wird am Computer durchgeführt. Das Programm wurde vom Kriminologischen Forschungsinstitut Niedersachsen e.V. erstellt. Dazu gehören Instruktion, zwei Fallbeschreibungen (Vignetten Betrug und Geldraub), Fragebogenitems zur Beurteilung der beschriebenen Delikte, Tappingprogramm und eine Mikrofonattrappe.

Zur Instruktion gehören Angaben über die Tastenbenützung, die Aufgabenstellung, Dank, Bitte um Stillschweigen über den Experimentvorgang, sowie Aufklärung über den Versuchshintergrund. Die Fallbeschreibungen sind Strafprozessinhalte, juristisch-formal gestaltet mit einem Block „Zum Sachverhalt“ und einem „Zu Täter und Opfer“. Die zwei Vignetten Betrug und Geldraub differenzieren sich in der emotionsauslösenden Inhaltskraft der Texte, variiert in der Information zum Täter. In der affektgeladenen Vignette ist der Täter vorbestraft, uneinsichtig und braucht das Geld für eigene Zwecke. Dahingegen in der Vignette mit niedrigem Affekt ist der Täter vorbestraft, Sozialhilfeempfänger, aber reuig und braucht das Geld prosozial, um für seinen kleinen Sohn sorgen zu können. Der Fragebogen besteht aus 14 Items zur Beurteilung des Schweregrades und der persönlichen Betroffenheit im Bezug auf das Delikt. Es bestehen fünf

Antwortmöglichkeiten von nicht bis stark betroffen. Darauf folgt eine letzte Frage zur Einschätzung eines bereits gegebenen Strafmasses von zwei Jahren Gefängnis unbedingt. Es bestehen wiederum 11 Antwortmöglichkeiten (-5 bis +5) von „viel zu gering“ über „angemessen“ bis „viel zu hoch“. Das Tappingprogramm besteht aus einem Signalton, der einen Rhythmus angibt, die Leerschlag Taste zu drücken. Dies dient der Ablenkung der Vpn in einer Versuchsbedingung. Die kognitive Verarbeitungskapazität sollte durch das Tappen absorbiert werden, der emotionale Aspekt der Urteilsbildung würde dadurch weniger unterdrückbar sein und das subjektive Strafurteil verstärken. Das Mikrofon täuscht das Aufnehmen des Vorlesens vor.

Durchführung

Die Versuche werden einzeln in einem Arbeitsraum der Universitätsbibliothek Tobler durchgeführt. Die Experimentinstruktion sowie abschliessende Erklärungen der Hintergründe sind im Computerprogramm enthalten. Die Versuchsleitung begrüsst die Vpn, erfragt die Daten für das Versuchspersonenprotokoll und führt die postexperimentelle Befragung durch. Zu allen vier Versuchsbedingungen gehört zuerst ein viermaliges Übungstapping à 40 sec. Nun erscheint die erste Fallbeschreibung eines Strafverfahrens, ein Betrugsdelikt. Je nach Versuchsbedingung wird die Vp aufgefordert, während des Vorlesens des Textes in das

Mikrofon, mitzutappen (sie beginnt damit schon kurz bevor der Text erscheint). Die andere Hälfte der Vpn liest nur ins Mikrofon. Die Fallbeschreibungen sind wiederum gleichmässig aufgeteilt in affektiv hohen oder niedrigen Gehalt.

Das 2x2 Design besteht aus zwei unabhängigen Variablen (UV) mit je zwei Bedingungen, UV1 ist der Affektgehalt der zu beurteilenden Straftat (niedriger Affekt vs hoher Affekt), UV2 die Ablenkung mit der Tappingbedingung (mit vs ohne Tapping). Als abhängige Variable (AV) gilt das Strafbeurteilungsmass, erfasst durch die Fragebögen in Interaktion mit der Tappingbedingung.

Nach dem Lesen beurteilt die Vp die Straftat anhand 14 Aussagen auf einer Skala mit fünf Antwortmöglichkeiten von „stimmt überhaupt nicht“ bis zu „stimmt voll und ganz“. Zum Schluss wird ein vorgegebenes Strafmass beurteilt. Dies auf einer Elferskala (-5 bis +5) von „viel zu gering“, über „angemessen“ bis „viel zu hoch“. Nun folgt eine zweite Fallbeschreibung (Raub) mit anschliessender Bewertung der gleichen Items. Nur diese Antworten sind schlussendlich für die Auswertung relevant. Zudem erfolgt postexperimentell ein Recalltest zur Überprüfung des inhaltlichen Verständnisses, ob die Vpn trotz Ablenkung den Inhalt der Geschichte erfasst haben. Zur Übersicht des Experimentablaufs dient Tabelle 1.

((Tabelle 1))

Resultate

Unsere Hypothese lautet: Vpn, die bei der Beurteilung normabweichenden Verhaltens durch eine gleichzeitig auszuführende Doppelaufgabe daran gehindert werden, ihre Gefühle zu kontrollieren, fordern härtere Strafen als Vpn, die nur mit der Beurteilung der Straftaten beschäftigt sind.

Nach Eingabe der Daten in die SPSS Version 10.0 wurde eine zweifaktorielle Varianzanalyse, univariat einseitig gerichtet, mit festen Faktoren „Tapping“ und „Affekt“ gerechnet. Als abhängige Variable wurde die Angemessenheit des Strafmasses erhoben.

((Tabelle 2))

Wie die Abbildung 1 zeigt (Abb.1), liegt der Mittelwert bei der Bedingung ohne Tapping bei $M = 1.64$. Der Mittelwert bei der Bedingung mit Tapping beträgt $M = 2.26$.

Es gibt keinen signifikanten Unterschied mit und ohne Tapping in der Beurteilung des Strafmasses $F(1,60) = 1.856, p = .178$ (Tabelle 2). In beiden Bedingungen wird das Strafmass als zu hoch eingeschätzt. Die affekt hohe Bedingung unterscheidet sich hoch signifikant von der affekt niedrigen Bedingung in der Beurteilung des Strafmasses. $F(1,60) = 13.701, p < .001$. Entgegen unserer Erwartung zeigte die zweite Hypothese einen gegenläufigen Effekt derart, dass die Vpn mit der Tapping-Bedingung eher das Gefühl hatten, dass die Strafe zu hoch ist, als diejenigen Vpn ohne Tapping Bedingung, was gegen unsere

Hypothese spricht. In der affekt niedrigen Bedingung wie auch in der affekt hoher Bedingung wird das Strafmass als zu hoch beurteilt. In der affekt niedrigen Bedingung ($M = 2.72$) ist die Beurteilung signifikant höher als in der affekt hohen ($M = 1.16$).

((Tabelle 3))

Die Interaktion, derart, dass der Mittelwert hoch/mit Tapping grösser ist als die übrigen drei Bedingungen, war wider Erwarten nicht signifikant. $F(1,60) = .753$, $p = .389$.

Die nachfolgende Auswertung der Resultate bezieht sich nicht auf aus der Theorie abgeleitete Hypothesen, sondern ist vielmehr ein Manipulation Check. Unsere erste Hypothese des Postexperimentes lautet, dass Vpn mit Tapping gleichviele Items wie Vpn ohne Tapping nennen.

Nach Eingabe der Daten in die SPSS Version 10.0 wurde ein t-Test für unabhängige Stichproben mit einem Alphafehler von 0.20 gerechnet. Der t-Test für unabhängige Stichproben zeigt, dass es keinen signifikanten Unterschied zwischen dem Mittelwert der Itemzahl genannt mit Tapping ($M = 9.68$) und dem Mittelwert der Itemzahl genannt ohne Tapping ($M = 9.94$) gibt; $t(32) = .501$, $p = .618$. Die erste Hypothese, dass Vpn mit Tapping gleichviele Items wie Vpn ohne Tapping nennen, hat sich bewährt.

Die zweite Hypothese des Postexperimentes lautet: Vpn mit der affekt hohen Geschichte nennen ein anderes Tätermotiv

(Geld für DVD-Player → Eigennutz) als Vpn mit der affekt niedrigen Geschichte (Geld für Sohn → prosozial).

Mit der zweiten Hypothese wollten wir überprüfen, ob diejenigen Vpn, welche die affekt hohe Geschichte gelesen haben, das Tätermotiv als eigennützig einschätzten und gleichzeitig diejenigen, welche die affekt niedrige Geschichte gelesen haben, das Tätermotiv als prosozial einschätzten. 100 % der Vpn haben in der affekt hohen Bedingung das Tätermotiv aus Eigennutz und in der affekt niedrigen Bedingung das Tätermotiv prosozial erkannt. $\chi^2(1, N = 63) = 63, p < .001$. Eine Vp konnte nicht zugeordnet werden. Unsere Hypothese, dass Vpn mit einer affekt hohen Geschichte ein anderes Tätermotiv nennen als Vpn mit einer affekt niedrigen Geschichte hat sich bestätigt.

Diskussion

Versuchspersonen, die bei der Beurteilung normabweichenden Verhaltens durch eine gleichzeitig auszuführende Doppelaufgabe daran gehindert werden, ihre Gefühle zu kontrollieren, fordern härtere Strafen als Versuchspersonen, die nur mit der Beurteilung der Straftaten beschäftigt sind.

Die Versuchspersonen der Doppelaufgaben-Bedingung fordern nicht signifikant härtere Strafen als Versuchspersonen, die nur mit der Beurteilung der Fälle beschäftigt sind. Der Haupteffekt Affekt wird hoch signifikant im Gegensatz zum Haupteffekt Tapping sowie der Interaktion.

Bei der Urteilsbedingung mit hohem Affektgehalt wird das Urteil als angemessener empfunden, gemessen an der Bedingung mit niedrigem Affektgehalt, unabhängig ob getappt wird oder nicht. Die Versuchspersonen mit der Tapping-Bedingung hatten eher das Gefühl, dass die Strafe zu hoch ist, als jene ohne Tapping-Bedingung, was gegen unsere Hypothese spricht. Mögliche Gründe dafür sind Störvariablen beim gesamten Experimentaufbau, der Versuchspersonen, der Versuchsleitung und unbekannte externe Einflüsse während der Durchführung des Experiments. Es könnte aber auch sein, dass die Theorie per se falsch ist.

Experimentaufbau: Wir konnten nachweisen, dass allfällige Unterschiede im Urteil nicht durch unterschiedliche Verständnisse des Textes zustande kamen. Durch unsere postexperimentelle, offene Befragung mittels eines Fragebogens wurde ersichtlich, dass der Inhalt der Geschichte auch über die Tapping-Bedingung richtig erfasst und somit verstanden wurde. Das heisst, dass Versuchspersonen mit der Tapping-Bedingung gleichviele Items post-hoc nennen konnten wie Versuchspersonen ohne die Tapping-Bedingung. Weiter wurde auch bei der „Affekt-hoch“-Geschichte sowie bei der „Affekt-niedrig“-Geschichte mit hundertprozentiger Sicherheit das richtige Tätermotiv, Eigennutz vs. prosoziales Verhalten, genannt.

Der kognitive Aufwand, der für das Tapping benötigt und der Affektkontrolle entzogen wird, genügt nicht, um den affektgeladenen Handlungsimpuls auszulösen. Das heisst, dass zu wenig kognitive Ressourcen entzogen werden während oder kurz vor der Beurteilung der Geschichten durch entweder lautes Vorlesen der Geschichte, Tapping oder beides. Unsere Vorschläge hierzu sind, einen unregelmässigen Tapping-Rhythmus anzuschlagen, den man nicht nach mehreren Durchgängen automatisieren kann, mit „schwächerer“ Hand tappen oder den Versuchspersonen die Instruktion zu einem möglichst klar artikulierten Vorlesungsstil vor dem Experiment vorzulegen.

Es könnte auch in Erwägung gezogen werden, Fragen zum Tapping mit ein zu beziehen, da das Tapping nach der ersten Geschichte nicht mehr gleich ernst genommen wird, da klar wird, dass die Geschichten wichtiger sind. Auch besteht ein möglicher unbekannter Einfluss der ersten Übungsvignette auf die zweite, d.h. eventuell wurden die Geschichten eher miteinander verglichen als unabhängig voneinander bewertet.

Der Einfluss der gewählten Sinnesmodalität (motorische Ablenkung vs. akustische Ablenkung etc.) muss abgeklärt werden. Das zu beurteilende Strafmass von 2 Jahren ist inadäquat zur Strafhandlung, d.h., 2 Jahre wird allgemein als zu hoch angesehen. Dies kann zu allgemeinen Urteilsverzerrungen führen. Das Strafmass einer unbedingten Freiheitsstrafe für einen Raub sollte reduziert werden.

Versuchspersonen: Es besteht auch ein möglicher Einfluss von musikalischen Fertigkeiten (Taktgefühl), die die Tapping-Leistung verbessern könnte. Die Auswahl der Versuchspersonen (Psychologie-studentinnen und -studenten), die Vorlesungsangst (Angst vor Blossstellung) und die momentane Stimmung, Motivation oder der Wettereinfluss können auf das Resultat einen Einfluss haben.

Versuchsleitung: Mögliche Beeinflussung, verbal wie auch nonverbal durch die Präsenz der Versuchsleitung während des Versuchs kann nicht ausgeschlossen werden. Da unsere Hypothese auf wissenschaftlichen Befunden basiert und ein Trend in diese Richtung bereits mehrfach erwiesen werden konnte, bleibt für unser Experiment die Frage offen, welche Störvariablen zur Falsifizierung unserer Hypothese beitragen.

Warum besteht die Annahme, dass Menschen Straftaten möglichst rational beurteilen und Emotionen unterdrücken? Kann man dies auch auf unsere Vignetten generalisieren? Fühlen sich die Versuchspersonen für ein gerechtes Urteil überhaupt verantwortlich oder ist der Aufwand für differenziertes und emotionales rationales Denken zu gross?

Schwer eruierbar ist auch, welche Emotionen die Geschichten bei den Versuchspersonen ausgelöst haben. Haben sie überhaupt Emotionen ausgelöst? Vielleicht sollten die Straftaten in den Geschichten brutaler sein, so dass man fast nicht emotional davon unberührt bleiben kann.

Schwer validierbar bleibt der gerechte Wert des vorgegebenen Strafmasses für die Tat der Vignetten (deutsches Recht vs. schweizerisches Recht). Zudem lassen die geringe Zahl von Versuchspersonen einen weiteren Test der Vignetten nötig erscheinen.

Autoren Anmerkungen

Corinne Gröbli, Studentin am Institut für Psychologie, Universität Bern; Sibylle Hafner, Studentin am Institut für Psychologie, Universität Bern; Marianne Spycher Studentin am Institut für Psychologie, Universität Bern; Bruno Sternath Student am Institut für Psychologie, Universität Bern; Antonia Wolleb Studentin am Institut für Psychologie, Universität Bern.

Wir danken Dr. Ute Gabriel für Ihre wertvollen Kommentare und Unterstützung.

Kommentare zu diesem Bericht sollten an die folgende Adresse gesendet werden, Bruno Sternath, Hombergstrasse 65, 4600 Olten Schweiz. E-Mail: bruno.sternath@bluewin.ch

References

- Applegate, B. K., Cullen, F. T., Link, B. G., Richards, P. J., Lanza-Kaduce, L. (1996). Determinants of public punitiveness toward drunk driving: A factorial survey approach. *Justice Q*, 13, 57-79.
- Eagly, A. H. Chaiken, S. (1993). *The psychology of attitudes*. Fort Worth, TX: Harcourt Brace Jovanovich.
- Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, 108, 814-834.
- Jacoby, J. E., Cullen, F. T. (1999). The structure of punishment norms: applying the Rossi-Beck model. *J Crim Law Crim*, 89, 245-312.
- Löbmann, R., Suhling, S., Greve, W, (2003). *The Emotional side of punitiveness: Affectively and cognitively based penal responses as a function of response mode*. Manuscript, submitted for publication.
- McKelvie, S. J., Mitchell, M., Arnott, R., Sullivan, M. (1993). Effects of offenders' and victims' characteristics on severity of punishment. *Psychological Rep.*, 72, 399-402.
- Miller, J. L., Rossi, P. H., Simpson, J. E. (1986). Perception of justice: Race and gender differences in judgments of appropriate prison sentences. *Law Social Review*, 20, 313-334.

Petty, R. E. Cacioppo, J. T. (1986). *Communication and persuasion: Central and peripheral routes to attitude change*. New York: Springer.

Tabelle 1: Experimentalablauf

E1 (Experimentalgruppe 1), E2 (Experimentalgruppe 2), K1 (Kontrollgruppe 1), K2 (Kontrollgruppe 2), Bedingung 1 (mit Tapping/affektgeladen), Bedingung 2 (mit Tapping/affektlos), Bedingung 3 (ohne Tapping/affektgeladen), Bedingung 4 (ohne Tapping/affektlos) t1 bis t7 (zeitlicher Ablauf)

	t1	t2	t3	t4	t5	t6	t7	t7
E1	Daten über Vp	vier-maliges Uebungs-tappen	Bericht vorlesen mit Bedingung 1 Uebung	Fragebogen zur Straftat-bewertung	Bericht vorlesen mit Bedingung 1	Fragebogen zur Straftat-bewertung	Recalltest	Postexperimentelle Befragung
E2	Daten über Vp	vier-maliges Uebungs-tappen	Bericht vorlesen mit Bedingung 2 Uebung	Fragebogen zur Straftat-bewertung	Bericht vorlesen mit Bedingung 2	Fragebogen zur Straftat-bewertung	Recalltest	Postexperimentelle Befragung
K1	Daten über Vp	vier-maliges Uebungs-tappen	Bericht vorlesen mit Bedingung3 Uebung	Fragebogen zur Straftat-bewertung	Bericht vorlesen mit Bedingung 3	Fragebogen zur Straftat-bewertung	Recalltest	Postexperimentelle Befragung
K2	Daten über Vp	vier-maliges Uebungs-tappen	Bericht vorlesen mit Bedingung4 Uebung	Fragebogen zur Straftat-bewertung	Bericht vorlesen mit Bedingung 4	Fragebogen zur Straftat-bewertung	Recalltest	Postexperimentelle Befragung

Tabelle 2:

Mittelwerte Bewertung der Straftat

		<i>M</i>	<i>SD</i>	<i>N</i>
Ohne Tapping	Affekt hoch	2.25	1.91	16
	Affekt niedrig	1.06	1.43	17
	Total	1.64	1.76	33
Mit Tapping	Affekt hoch	3.19	1.68	16
	Affekt niedrig	1.27	1.67	15
	Total	2.26	1.91	31
Total	Affekt hoch	2.72	1.84	32
	Affekt niedrig	1.16	1.53	32
	Total	1.94	1.85	64

Tabelle 3

Ergebnistabelle der zweifaktoriellen Varianzanalyse

Q.d.V	QS	df	F	p
Tapping	5.237	1	1.856	.178
Affekt	38.663	1	13.701	.000
Tapping x Affekt	2.125	1	.753	.389
Fehler	169.312	60		
Total	456.000	64		

Abbildung 1. Mittelwerte der 2-faktoriellen Varianzanalyse.

