

Barton/Dubelaar/Kölbel/Lindemann (Hrsg.)

„Vom hochgemuten, voreiligen Griff nach der Wahrheit“

Fehlurteile im Strafprozess





Interdisziplinäre Studien zu Recht und Staat

begründet von Werner Maihofer[†] und Gerhard Sprenger[†]

In Verbindung mit

Dieter Grimm, Eric Hilgendorf, Joachim Hruschka[†],
Hermann Klenner, Ernst-Joachim Lampe, Stefan Oeter,
Britta Padberg, Manfred Rehbinder, Hubert Rottleuthner,
Rüdiger Schott[†]

herausgegeben von

Tatjana Hörnle, Stefan Huster und Gertrude Lübbe-Wolff

Band 56

Stephan Barton/ Marieke Dubelaar/Ralf Kölbel
Michael Lindemann (Hrsg.)

„Vom hochgemuten, voreiligen Griff nach der Wahrheit“

Fehlurteile im Strafprozess



Nomos

Die Interdisziplinären Studien zu Recht und Staat sind eine neue Folge des Jahrbuchs für Rechtssoziologie und Rechtstheorie, das in den Jahren 1970–1993 beim Bertelsmann Universitätsverlag Reinhard Mohn bzw. Westdeutschen Verlag erschienen ist.

Bereits in den letzten Jahren gingen die Inhalte sachlichen Notwendigkeiten gehorchend zunehmend über die Bereiche Rechtssoziologie und Rechtstheorie hinaus. Fast immer wurden auch grundlegende Fragen der Rechtsphilosophie und Staatstheorie mitbehandelt. Dies soll im neuen Titel der Reihe zum Ausdruck kommen und noch ein anderes: die in den Jahrbuch-Bänden veröffentlichten Beiträge und Diskussionen haben durchweg fachübergreifenden Charakter. Sie sind überwiegend aus Tagungen hervorgegangen, die durch das *Zentrum für interdisziplinäre Forschung (ZiF)* der Universität Bielefeld gefördert wurden. Auch dies sollte in dem neuen Namen deutlich werden.

Überlingen/Bielefeld 1994

Die Herausgeber des Jahres 1994



Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

ISBN 978-3-8487-4891-4 (Print)

ISBN 978-3-8452-9080-5 (ePDF)

1. Auflage 2018

© Nomos Verlagsgesellschaft, Baden-Baden 2018. Gedruckt in Deutschland. Alle Rechte, auch die des Nachdrucks von Auszügen, der fotomechanischen Wiedergabe und der Übersetzung, vorbehalten. Gedruckt auf alterungsbeständigem Papier.

Vorwort

Am 23. und 24. November 2017 fanden die 6. Bielefelder Verfahrenstage im Zentrum für interdisziplinäre Forschung der Universität Bielefeld statt. Wie schon bei den Vorgängertagungen stand auch diesmal neben dem interdisziplinären Dialog vor allem der Austausch zwischen Wissenschaft und Praxis im Vordergrund. Eine Neuerung bestand darin, dass in deutlich stärkerem Maße als in der Vergangenheit auch Forscherinnen und Forscher aus dem Ausland an der Durchführung der Tagung beteiligt waren: Mit Marieke Dubelaar (Universität Leiden, NL) wirkte erstmals eine ausländische Wissenschaftlerin an der Organisation der Tagung mit, und zwei der vier Arbeitsgruppen wurden in englischer Sprache abgehalten, um die Rezeption des internationalen Forschungsstandes zu ermöglichen. Das Experiment einer behutsamen Internationalisierung des bewährten Formates ist nach Ansicht der Veranstalter durchaus gelungen.

Für die Durchführung der Tagung konnten wir erneut auf die tatkräftige Mitwirkung engagierter Referenten und Moderatoren sowie zahlreicher im Hintergrund wirkender helfender Hände zurückgreifen. Danken möchten wir Silke Hüls für die Übernahme der Moderation einer Arbeitsgruppe und Thomas Fischer, Jens Gnisa, Stefan König und Sabine Rückert für die Mitwirkung an der abschließenden Podiumsdiskussion. Christa Kleinhagenbrock danken wir für die sorgfältige Transkription der Diskussion. Dank gebührt auch den Teams der beiden Bielefelder Lehrstühle für die unermüdliche Unterstützung bei der Organisation und Durchführung der Tagung sowie Frau Monika Faßbender vom Münchener Lehrstuhl für Strafrecht und Kriminologie für die Mitwirkung bei der Edition dieses Tagungsbandes.

Bedanken möchten wir uns schließlich bei der Juristischen Gesellschaft Ostwestfalen-Lippe, die einen großzügigen Tagungszuschuss gewährt hat und auf diese Weise nicht nur zum Gelingen der Tagung beigetragen, sondern auch das Erscheinen dieser Publikation ermöglicht hat.

Bielefeld, Leiden und München im Mai 2018

Stephan Barton, Marieke Dubelaar, Ralf Kölbel und Michael Lindemann

Inhaltsverzeichnis

<i>Stephan Barton, Marieke Dubelaar, Ralf Kölbl, Michael Lindemann</i> Aspekte der Fehlurteilsforschung: Zur Einführung in den Band	9
<i>Ralf Kölbl</i> Der Stand der internationalen Fehlurteilsforschung: Was kann man daraus für Deutschland lernen?	31
<i>Gwladys Gilliéron</i> Fallstricke für die Wahrheitsfindung in summarischen Verfahren	59
<i>Jörg Kinzig</i> Freispruch nach Untersuchungshaft – Folgerungen für die Fehlurteilsforschung	79
<i>Margit E. Oswald, Helen Wylér</i> Fallstricke auf dem Weg zur »richtigen« Entscheidung im Strafrecht: Eine Analyse aus psychologischer Sicht	103
<i>Anna Sagana</i> The downward spiral of biases in criminal investigations: From eyewitnesses to forensic experts and judges	133
<i>Marieke Dubelaar</i> Use of scenarios in evidential reasoning from a continental law perspective	147
<i>Dave van Toor</i> Will neuroscience revolutionize criminal interrogations?	179
<i>Stephan Barton</i> Benötigen wir psychologische Sachverständige für die Glaubhaftigkeitsbeurteilung?	199

<i>Karsten Altenhain</i>	
Vor- und Nachteile der audiovisuellen Aufzeichnung von Zeugenaussagen	225
<i>Thomas Fischer</i>	
Warum lässt das Revisionsrecht Fehlerurteile zu?	253
<i>Cyrille Fijnaut</i>	
Reviewing criminal cases from the perspective of comparative law	275
Verzeichnis der Autorinnen und Autoren	303

Fallstricke auf dem Weg zur »richtigen« Entscheidung im Strafrecht: Eine Analyse aus psychologischer Sicht

Margit E. Oswald, Helen Wylter

A. Einführung

Fragen wir uns zu Beginn, warum sich Strafrechtswissenschaftler und Strafrichter¹ in Deutschland erst jetzt intensiver für Urteilsfehler zu interessieren beginnen, über die in der Psychologie bereits seit den 70er Jahren des letzten Jahrhunderts intensiv geforscht wird. Für den deutschsprachigen Raum sind uns zumindest nur wenige rechtswissenschaftlich ausgerichtete Tagungen bekannt, die sich speziell mit dieser Thematik auseinandergesetzt haben. Eine Ausnahme bildet wohl die im Mai 2004 an der FU Berlin stattgefundene 94. Dahlem Konferenz zum Thema »Heuristics and the Law« (*Gigerenzer & Engel 2006*), die sich aber nicht gezielt auf das Strafrecht bezog.

Ein Grund für das relativ spät aufkommende Interesse von Juristen an psychologischen Erkenntnissen könnte sein, dass die grosse Anzahl an Publikationen, die sich thematisch mit Urteilsfehlern auseinandersetzt, nach den vielen Jahren kaum noch übersehen werden kann. Ein zweiter Grund könnte aber auch sein, dass in der Psychologie sowohl aus theoretischer wie empirischer Sicht zunehmend speziell auf strafrichterliche Entscheidungen Bezug genommen wird. Der vielleicht wichtigste Grund wird aber sein, dass Juristen spätestens aufgrund der Erkenntnisse des in den USA angesiedelten »Innocence Project« (*Drizin & Leo 2004*) das Problem von Fehlurteilen im Strafrecht nicht mehr ignorieren können.

In der Sozialpsychologie gab es bereits in den 60er Jahren wichtige Erkenntnisse über Fehler, die Personen bei der Beurteilung ihrer eigenen

1 Aufgrund der besseren Lesbarkeit wird in diesem Text weitgehend auf die explizite zusätzliche Erwähnung der weiblichen Form verzichtet. Wir weisen an dieser Stelle ausdrücklich darauf hin, dass die männliche Form immer als die weibliche Form miteinschliessend zu verstehen ist.

oder einer anderen Person begehen, wie z.B. der Fundamentale Attributionsfehler, auch »correspondence bias« genannt. Jones & Harris konnten 1967 in einem Experiment zeigen, dass Personen die in der Umwelt befindlichen Verhaltensursachen stark vernachlässigen und zu schnell vom beobachteten Verhalten auf eine dahinterliegende *Persönlichkeitseigenschaft* schliessen. Als Meilenstein der Fehlerforschung ist aber gleichwohl der 1974 erschienene Aufsatz von Amos Tversky & Daniel Kahneman zu sehen. Die Autoren zeigen, dass Menschen beim Denken einfache Faustregeln, sogenannte Heuristiken, verwenden, die zwar nicht zwangsläufig, aber doch gelegentlich zu Fehlern oder Verzerrungen (Biases) führen. Bekannt sind vor allem der Repräsentativitäts- und der Verfügbarkeitsheurismus. Der Repräsentativitätsheurismus besagt in knappen Worten, dass sich Personen bei dem Urteil, ob ein konkretes Ereignis (z.B. Jugendlicher schaut *gewalthaltige* Videos) diagnostisch relevant ist für eine bestimmte Kategorie (z.B. *Gewaltverbrechen*), von *Ähnlichkeiten* leiten lassen und dabei die Basisrate des Ereignisses (z.B. Häufigkeit, mit der Jugendliche im Allgemeinen gewalthaltige Videos schauen) vernachlässigen. Sobald aber die Basisrate des Ereignisses relativ hoch ist, kommt es durch ihre Vernachlässigung zu einer systematischen Überschätzung der diagnostischen Relevanz des konkreten Ereignisses. Der Verfügbarkeitsheurismus bezieht sich auf die Einschätzung der Häufigkeit eines Ereignisses. Sind Einzelbeispiele eines Ereignisses »schnell« aus dem Gedächtnis abrufbar, so schätzen wir deren Auftretenshäufigkeit höher ein, als wenn uns der Abruf aus dem Gedächtnis schwerfällt. Dies kann zu Fehlern führen, wie Tversky & Kahneman (1974) anhand mehrerer Beispiele zeigen. So meinen wir beispielsweise irrtümlich, dass es mehr Worte mit drei oder mehr Buchstaben gibt, die ein »r« am Anfang haben, als Worte, die ein »r« an dritter Stelle aufweisen.

Zeitlich etwas später postulieren Kahneman & Tversky in ihrer Prospekttheorie (Kahneman & Tversky 1979) den Framing-Effekt, der besagt, dass sich Personen bei der Wahl zwischen zwei Alternativen mit identischem Erwartungswert unterschiedlich verhalten, je nachdem, ob man die Alternativen in einem Gewinnrahmen (gain frame) oder aber einem Verlustrahmen (lost frame) *darstellt*. Dies widerspricht dem Rationalitätskriterium der »Invarianz«, wonach sich Personen *unabhängig vom Kontext*

jeweils für die Alternative mit dem größten Nutzen (Wert x Erwartung)² entscheiden sollten. Wird die eine Wahl als Alternative zwischen möglichen Verlusten wahrgenommen, so sind Personen risikogeneigt (wählen nicht das sichere, sondern das unsichere Ereignis). Wird die Wahl hingegen als Alternative zwischen möglichen Gewinnen wahrgenommen, so verhalten sie sich risikoscheu (wählen das sichere Ereignis). So wählen bei Gericht die Beklagten, die am Gewinnen sind, eher den Vergleich (= sicheres Ereignis) als einen unsicheren Gewinn (= den Prozess weiterziehen), sind also risikoscheu. Kläger hingegen, die am Verlieren sind, entscheiden sich eher für das unsichere Ereignis (= Prozess weiterziehen) als für den vorgeschlagenen Vergleich (= sicherer Verlust), sind also risikogeneigt (vgl. *Rachlinski* 2000). Die tatsächlichen Gerichtsdaten stimmen übrigens recht gut mit den Voraussagen der Prospekttheorie überein (vgl. *Schweizer* 2005).

In den folgenden Jahrzehnten vervielfachte sich die Anzahl aufgedeckter Urteilsfehler. So wird die »Kontrollillusion« entdeckt (*Langer* 1975), nach der Menschen die Tendenz haben, zu glauben, gewisse Vorgänge kontrollieren zu können, die nachweislich nicht beeinflussbar sind, wie z.B. Vorgänge beim Würfel- oder Münzwurfspiel. Des Weiteren wird das »Hot-Hand-Phänomen« entdeckt (*Gilovich, Vallone & Tversky* 1985), das sich auf den Irrglauben bezieht, bei voneinander unabhängigen Ereignissen eine *Glücksträhne* zu haben, wenn bereits mehrere Erfolge vorausgegangen sind, oder der »Bias Blind Spot« (*Pronin, Lin & Ross* 2002), wonach Personen die oft unbegründete Tendenz haben, sich für objektiv und unbeeinflusst zu halten.

Soweit der kleine Einblick in die Vielzahl der mittlerweile diskutierten Heuristiken und Urteilsfehler. Jedoch möchten wir die Leser nicht weiter verwirren, indem wir mit dieser Aufzählung von Heuristiken und Urteilsfehlern fortfahren, sondern das Kapitel an folgenden Fragen ausrichten: Welche bekannten Urteilsfehler spielen im Rahmen der Strafverfolgung eine besondere Rolle? Was wissen wir über diese Urteilsfehler und welche Erkenntnisse gibt es speziell über Urteile von Strafrichtern? Und können

2 Die Formel bezieht sich auf den von der Person subjektiv wahrgenommenen Wert eines bestimmten Ereignisses (Wert) und der erwarteten Wahrscheinlichkeit, dass dieses Ereignis tatsächlich eintritt (Erwartung).

diese Urteilsfehler vermieden werden, bzw. welche Trainingsstrategien sind aus psychologischer Sicht erfolgreich?

B. Drei Urteilsheuristiken, bzw. Urteilsfehler, die für das Strafrecht eine besondere Rolle spielen

Viele der Urteilsheuristiken und Urteilsfehler, wie beispielsweise der Repräsentativitätsheurismus oder der Framing-Effekt, sind für eine kritische Urteilsanalyse im Rahmen der Strafverfolgung zweifellos relevant (vgl. Guthrie, Rachlinski & Wistrich 2001; Schweizer 2005). Drei Urteilsfehler erscheinen uns aber von besonderem Interesse zu sein: die Positive Teststrategie, bzw. der Bestätigungsfehler, der Rückschaufehler und der Ankereffekt.

Diese drei Urteilsfehler sind in verschiedenen Phasen des richterlichen Urteilens prominent vertreten (vgl. Abbildung 1), auch wenn sie nicht ausschliesslich dort auftreten, und können jeweils hinreichender Grund für erfolgreiche Berufungen oder Wiederaufnahmeverfahren sein. Des Weiteren handelt es sich bei ihnen um besonders robuste Urteilsfehler, da sie in verschiedenen Kontexten und bei verschiedenen Personengruppen nachgewiesen werden können. Schließlich gibt es zu ihnen Untersuchungen, die sich auch auf die Gruppe der Strafrichter beziehen.

Unterscheiden wir zwischen drei Urteilsphasen, dem Schuldurteil, dem Urteil über die Subsumption der Straftat und der Festlegung von Art und Höhe der Strafe innerhalb des vorgegebenen Strafrahmens, so spielen die Positive Teststrategie, bzw. der Bestätigungsfehler vor allem bei der Schuldfrage, also bei der Frage, ob der Angeklagte die Tat begangen hat, eine besondere Rolle. Der Rückschaufehler spielt hingegen insbesondere bei der Subsumption der Tat eine Rolle und der Ankereffekt schliesslich bei der Höhe der Strafzumessung. Ausgespart bei dieser zugegebenermassen recht groben Dreiteilung des Urteilsprozesses bleiben beispielsweise die Entscheidung über die Schuldfähigkeit des Angeklagten oder das Hinzuziehen externer Gutachter.

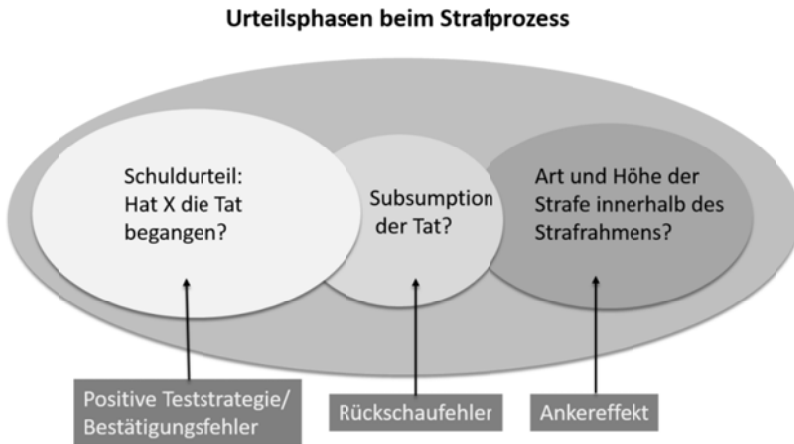


Abbildung 1: Zuordnung wichtiger Heuristiken und Urteilsfehler zu verschiedenen Phasen des strafrichterlichen Urteilens.

I. Positive Teststrategie und Bestätigungsfehler

Obwohl viele Autoren die Positive Teststrategie und den Bestätigungsfehler quasi als Synonyme verwenden³, bestehen doch bedeutsame Unterschiede zwischen ihnen, die vor allem von *Klayman & Ha* (1987) klar herausgearbeitet wurden.

Die Positive Teststrategie bezeichnet ein kaum vermeidbares Such- und Frageverhalten von Personen. Wann immer eine Annahme zu überprüfen ist, fragen/suchen Personen auf eine ganz bestimmte Weise nach Information. Möchten sie beispielsweise wissen, ob eine andere Person extravertiert ist, so fragen sie typischerweise »Sind Sie gerne unter Menschen?« oder »Gefällt es Ihnen, eine Party in Schwung zu bringen?« (vgl. *Snyder & Swann* 1978). Natürlich könnte man auch fragen, ob die Person gerne allein ist, oder ob sie gerne Bücher liest, aber diese Art von Fragen werden eher nicht gewählt. Machen wir die Such- und Fragestrategie an noch ei-

3 Auch in neueren Studien wird die Positive Teststrategie als Bestätigungsfehler interpretiert und dabei sogar *Klayman & Ha* (1987) zitiert (vgl. *Bahník & Strack* 2016).

nem anderen Beispiel deutlich. Sucht eine Person ihre Lesebrille, so sucht sie typischerweise an solchen Stellen, an denen sie meint, sie zuletzt getragen zu haben, oder aber wo sie die Brille oft ablegt. Wenig wahrscheinlich wird es hingegen sein, dass sie die Brille beispielsweise unterm Bett oder im Kühlschrank sucht. Eine Positive Teststrategie liegt also vor, *wenn Personen nach denjenigen kritischen Ereignissen fragen/suchen, die sie gemäß ihrer Annahme erwarten* (Klayman & Ha 1987, S. 212). Dass es zu dieser Strategie kommt, setzt noch nicht einmal eine Annahme oder Hypothese voraus, von der man persönlich überzeugt ist. Es reicht aus, dass Personen lediglich die Aufgabe erhalten, eine beliebige Annahme zu überprüfen.

Für die Abgrenzung vom Bestätigungsfehler ist es jedoch wichtig, festzuhalten, dass mit der Anwendung einer Positiven Teststrategie noch *keine Suche nach der Bestätigung* der Annahme verbunden ist. Bis auf sehr spezielle Ausnahmefälle (vgl. Wason 1960) erlaubt sie es nämlich, dass ein kritisches Ereignis *nicht* auftritt, die Annahme also falsifiziert wird, und dass die Person dies auch zur Kenntnis nimmt (vgl. Devine, Hirt & Gehrke 1990; Oswald & Grosjean 2004; Trope & Bassok 1982). So kann die befragte Person beim oben erwähnten Beispiel antworten, dass sie nicht gerne auf Partys geht, geschweige denn Gefallen daran findet, eine solche in Schwung zu bringen. Es handelt sich bei der Positiven Teststrategie also um ein Prüfverhalten, wodurch im Allgemeinen der *Suchraum auf eine sehr sinnvolle Weise eingeschränkt* wird.⁴ Natürlich ist es beispielsweise bei der Einvernahme von Beschuldigten wichtig, so lange als möglich zunächst *offene Fragen* zu stellen, die den Gesprächspartner weder in die eine noch in die andere Richtung einschränken (Vrij, Mann, Kristen & Fisher 2007), bevor man spezifische Fragen stellt, welche mit dem Prüfen einer oder mehrerer Annahmen des Befragers verbunden sind. Geht es also darum, die Wahrheit herauszufinden, so wird man gleichwohl nicht darum herumkommen, gegen Ende einer Befragung spezifische Fragen im Sinne der Positiven Teststrategie zu stellen.

4 Nach etwas zu suchen oder zu fragen, was man nicht erwartet (Negative Teststrategie) würde einen unbegrenzten Suchraum eröffnen und die Suche willkürlich gestalten, es sei denn, man hätte eine alternative Idee oder Annahme und sucht, bzw. fragt (auch) gemäss dieser Alternative. In diesem Fall würde man aber wiederum eine Positive Teststrategie anwenden.

Ein von der Positiven Teststrategie prinzipiell unabhängiges Problem ist darin zu sehen, dass Personen zur Erklärung eines Ereignisses oft nur *eine Annahme fokussieren* und nicht auch nach Alternativerklärungen suchen, oder aber diese ausblenden. Die Fokussierung einer einzelnen Annahme hat zur Folge, dass (a) einseitig, also nicht aus verschiedenen Blickwinkeln heraus spezifische Fragen im Sinne der Positiven Teststrategie gestellt werden, so dass es bei einer Bejahungstendenz der Befragten zu einer Urteilsverzerrungen kommen kann (Zuckerman, Knee, Hodgins & Miyake 1995) und es (b) schnell zu einer Überschätzung bestätigender Ereignisse kommt, speziell dann, wenn diese diagnostisch ambivalent sind (vgl. Kruglanski & Maysless 1988). Machen wir letztgenannte Überlegung an einem Beispiel deutlich: Sie sehen, wie ein Jugendlicher provoziert wird und den Stinkefinger zeigt (Ereignis X). Dieses Ereignis X ist erst dann diagnostisch relevant, wenn es nur oder aber deutlich häufiger bei aggressiven als bei nicht-aggressiven Jugendlichen auftritt.⁵ Haben Sie aber lediglich die Annahme im Fokus, dass es sich um einen aggressiven Jugendlichen handelt, so werden Sie den Stinkefinger möglicherweise selbst dann als Bestätigung Ihrer Annahme interpretieren, wenn das Ereignis diagnostisch nicht relevant ist. Die Fokussierung auf nur eine Annahme zur Erklärung eines Ereignisses führt uns daher unmittelbar zum Bestätigungsfehler.

Von einem Bestätigungsfehler kann ausgegangen werden, wenn Personen auf systematische Weise versuchen, ihre Annahme/Erwartung beizubehalten. In diesem Fall werden Informationen so gesucht oder interpretiert, dass eine Bestätigung der Annahme (unabhängig vom Wahrheitsgehalt) wahrscheinlich wird, und solche Informationen vermieden oder uminterpretiert, die der Annahme/Erwartung widersprechen (Koriat, Lichtenstein & Fischhoff 1980). Das systematische Ausblenden von Alternativerklärungen ist in dem Sinn also ein probates Mittel, um seine Annahmen/Erwartungen beizubehalten, weil auf diese Weise bestätigende Ereignisse in ihrer Bedeutung überschätzt werden. Bei einem Bestätigungsfeh-

5 Das Ereignis »Stinkefinger nach einer Provokation« (X) ist nur dann eine objektive Bestätigung der Annahme, dass der betreffende Jugendliche aggressiv ist, wenn $p(X/\text{Jugendlicher ist aggressiv})$ deutlich grösser ist als $p(X/\text{Jugendlicher ist nicht aggressiv})$. Ignoriere ich die Alternativerklärung, so verkenne ich evtl. die mangelnde Diagnostizität des Ereignisses.

ler werden aber nicht nur die bestätigenden Ereignisse in ihrer Bedeutung überschätzt, sondern vor allem auch die falsifizierenden Ereignisse umgedeutet oder gar ignoriert, wie oben erwähnt wurde. Während das Ausblenden von Alternativerklärungen noch als eine rein kognitive Strategie angesehen werden kann, ist der die eigene Annahme immunisierende Umgang mit erwartungskonträren Informationen nur noch schwer über ausschliesslich kognitive Prozesse zu erklären. Er dürfte vielmehr mit der Motivation verbunden sein, die Annahme/Erwartung *verteidigen* zu wollen, weil eine Widerlegung schmerzlich, oder zumindest mit negativen Emotionen verbunden wäre. Insgesamt dürfte der Bestätigungsfehler daher umso ausgeprägter sein, je mehr die Person (a) von ihrer anfangs bestehenden Annahme/Erwartung überzeugt und (b) je höher ein Erfolgs- und/oder Zeitdruck mit der Bestätigung der betreffenden Annahme verbunden ist (*Trope & Lieberman* 1996).

Vielleicht ist der Bestätigungsfehler nicht nur das Resultat einer einzelnen Heuristik, sondern mit dem Einsatz eines ganzen Bündels von Heuristiken verbunden, die umso eher Anwendung finden, je höher die Motivation ist, die betreffende Annahme/Erwartung zu verteidigen. Überlegenswert erscheint uns in diesem Zusammenhang daher die These, dass es sich bei dem Bestätigungsfehler um ein Kontinuum handelt, das vom Einsatz der Positiven Teststrategie bis hin zum Einsatz von Strategien der völligen Immunisierung reicht (vgl. Abbildung 2).

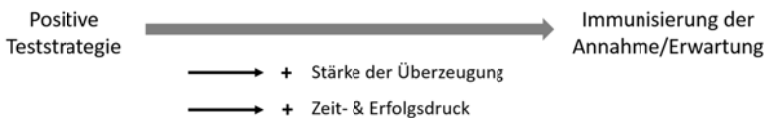


Abbildung 2: Postuliertes Kontinuum des Bestätigungsfehlers

Ist der Bestätigungsfehler eine Gefahr für faire Strafprozessentscheidungen? Grundsätzlich ist zu sagen, dass die Situation bei Gericht eine gewisse suggestive Potenz für einen Bestätigungsfehler aufweist. Die Person ist ja nicht umsonst angeklagt, und es scheint daher vielleicht »verführerisch« zu sein, vorschnell von ihrer Schuld auszugehen (*Rassin* 2017; *Volbert* 2013). Dies insbesondere dann, wenn auch der erste, oft im Bruchteil einer Sekunde gebildete Eindruck vom Angeklagten negativ ist und keine Alternativerklärungen für die Straftat auf der Hand liegen. Zudem besteht für Richter nicht selten Zeit- und Erfolgsdruck, und bei manchen Strafsachen auch ein medialer Druck, den Fall auf bestimmte Weise abzuschliessen.

In zahlreichen psychologischen Studien konnte belegt werden, dass eine *Schulderwartung*, die man im ersten Teil der jeweiligen Studie bei einer Gruppe der Probanden manipuliert hatte, den Urteilsprozess auf verschiedene Weise zu beeinflussen vermag:

(1) Wurde die Schulderwartung durch das Aussehen des Angeklagten verstärkt, indem das Foto z.B. einer *nicht* vertrauenswürdigen (vs. vertrauenswürdigen) Person gezeigt wurde, so reichte dies bereits aus, dass die Probanden *schneller zu einem vorläufigen Schuldurteil* kamen, bzw. für das Schuldurteil eine geringere Anzahl an belastenden Hinweisen benötigten als diejenigen, bei denen das Foto einen vertrauenswürdigen aussehenden Angeklagten zeigte (*Porter, ten Brinke & Gustaw* 2010).

(2) Wenn vom Angeklagten ein Geständnis vorliegt, so scheint dies nicht nur die Schulderwartung massiv zu fördern, sondern scheint auch die *Beurteilung des gesamten Beweismaterials zu beeinträchtigen* (vgl. *Kassin, Dror & Kukucka* 2013). So sahen die Teilnehmer einer Studie von *Kassin & Sukel* (1997) eine simulierte Gerichtsverhandlung, in der ein hoher Druck der Befragenden auf den Angeklagten ausgeübt wurde, wobei in der einen Gruppe der Angeklagte ein Geständnis abgab, in der anderen hingegen nicht. Obwohl die meisten Teilnehmenden korrekt erkannten, dass das Geständnis aufgrund des hohen Befragungsdrucks »kein zulässiges Beweismittel« ist, wurden dennoch mehr Schuldsprüche in der Bedingung mit Geständnis ausgesprochen als in der Bedingung ohne Geständnis. Dass bei gegebener Schulderwartung die Unabhängigkeit der Würdigung einzelner Beweismittel beeinträchtigt sein kann, konnten auch *Greenspan & Scurich* (2016) sowie *Ask & Granhag* (2005) zeigen.

(3) Eine Schulderwartung kann sich auch negativ auf die *Prüfgenauigkeit* bei einer forensischen Untersuchung auswirken, was in der Regel mit Folgen für das Richterurteil verbunden ist. *Kukucka & Kassin* (2014) konnten dies in Bezug auf die Fehlerhäufigkeit nachweisen, die beim Handschriftenvergleich auftritt. Sowohl *Dror, Charlton & Péron* (2006) als auch *Smalarz, Madon, Yang, Guyll & Buck* (2016) konnten einen ähnlichen Effekt bei der Beurteilung von Fingerabdrücken belegen. Letztere manipulierten die Schulderwartung der Teilnehmenden, indem Verdächtige entweder einem deliktsspezifischen Stereotyp entsprachen (z.B. weiss und männlich bei Kindsmisbrauch) oder aber nicht (z.B. asiatisch und weiblich bei Kindsmisbrauch). Beim Vergleich der Fingerabdrücke des Verdächtigen mit denjenigen auf dem Tatgegenstand wurden die objektiv verschiedenen Fingerabdrücke deutlich häufiger als identisch wahrgenommen.

nommen, wenn der Verdächtige dem deliktsspezifischen Stereotyp entsprach (vgl. Abbildung 3).

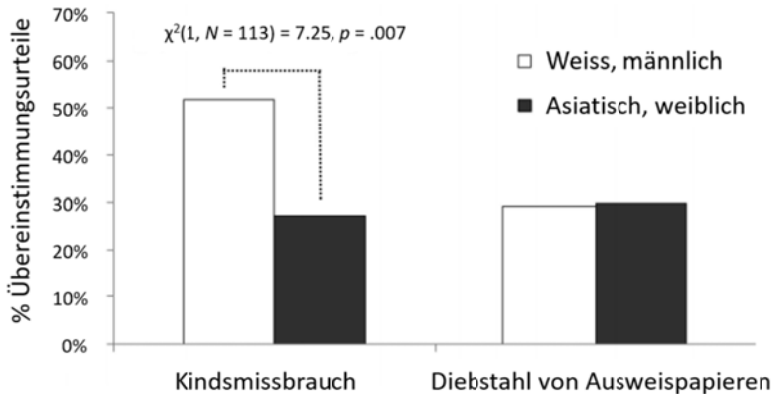


Abbildung 3: Einfluss von Delikttyp und Merkmale des Verdächtigen auf die Einschätzung, dass (ungleiche) Fingerabdrücke identisch sind (Smalarz et al., 2016, S. 425)

(4) Schließlich konnte gezeigt werden, dass bei höherer Schuldurwartung die verdächtige Person auch *konfrontativer befragt* wird. In einer mock-crime Studie von Narchet, Meissner & Russano (2011) werden die Teilnehmenden zu einem Betrug (Mogeln bei einer Examensarbeit) animiert. Je nachdem, ob der Animation Folge geleistet wurde oder nicht, gehörten die Teilnehmenden zur Gruppe der »Betrüger« oder aber zur Gruppe der »Unschuldigen«. Beide Teilnehmergruppen wurden anschliessend einvernommen. Erhielten die Befrager die zusätzliche Vorinformation, dass eine Videoaufnahme Anhaltspunkte für die Schuld der betreffenden Person liefert (Schuldmanipulation), so wurde deutlich konfrontativer befragt (Minimierungs- und Maximierungsstrategien) als wenn diese Vorinformation nicht gegeben wurde. Außerdem stieg die Anzahl falscher Geständnisse. Wurde keine Vorinformation gegeben, so legten Unschuldige im Schnitt in 20 % der Fälle ein falsches Geständnis ab, während sie in 47 % der Fälle ein falsches Geständnis abgaben, wenn die schulderhöhende Vorinformation gegeben und entsprechend konfrontativer befragt wurde.

Einschränkend ist allerdings zu sagen, dass fast alle oben genannten Studien an Studierenden und in Ausnahmefällen auch an Personen der Normalbevölkerung (vgl. Greenspan & Scurich 2016) durchgeführt wur-

den. Daher stellt sich sicher die berechnigte Frage, welche Folgerungen mit diesen Erkenntnissen für die Praxis von Strafrichtern verbunden sind. Sind sie besser als Studierende oder »normale« Laien, evtl. sogar gefeit gegenüber dem Bestätigungsfehler? Anlass zu einer solchen Vermutung gibt vor allem der Umstand, dass es für Richter mehrere protektive Faktoren gibt, die den Einfluss des Bestätigungsfehlers minimieren könnten. Strafrichter verfügen nicht nur über Expertenwissen, sondern sind auch verantwortlich für die Korrektheit ihrer Urteile, was in der Regel eine intensivere und objektivere Informationsverarbeitung zur Folge hat (vgl. *Schmittat & English* 2016). Außerdem sind sie der Konkurrenz der Meinungen ausgesetzt, wie z.B. den Meinungen der Staatsanwaltschaft und der Verteidigung, und alle Parteien haben die Möglichkeit an der Hand, Rechtsmittel gegen ein als unfair empfundenes Strafurteil einzulegen.

Weiterhin könnte eingewendet werden, dass es selbst für den Fall, dass einige oder mehrere Richter und Richterinnen nicht dagegen gefeit sind, einen Bestätigungsfehler zu begehen, sich dieser ja nicht unbedingt gravierend bemerkbar machen muss. In einer persönlichen Kommunikation vom 17. Oktober 2017 sagt der Präsident des Obergerichts Bern (CH), Herr *Stephan Stucki*: »Es gibt natürlich fast laufend Urteile, bei denen wir (Berufungskammer, zweite Instanz) die Beweiswürdigung eines erstinstanzlichen Gerichts korrigieren, seltener, dass wir das massiv tun. Dass man beweismäßig anders gewichten kann, kommt - wie gesagt - öfter vor, heißt ja noch nicht grad, dass ein eigentliches Fehlurteil vorliegt.«

Andererseits gibt es in den öffentlichen Medien aber immer wieder einmal Behauptungen, wie »Bei Gericht werden die Indizien oft so zusammengefügt, dass sie ins Bild passen« (*Janisch* 2017), und neben protektiven Faktoren gibt es klar auch *Risikofaktoren* für einen Bestätigungsfehler bei Gericht. So gibt es nachweislich Vorurteile bei Strafrichtern (vgl. *Oswald* 2014), die auch unbewusst einen Bestätigungsfehler in Gang setzen können. Vor allem aber dürften Zeit-, Erfolgsdruck und nicht zuletzt auch ein medialer Druck dem Bestätigungsfehler Nahrung geben. Als Beispiel soll hier auf den Fall Ignaz Walker, einen Indizienprozess des Urner Obergerichts (CH), verwiesen werden, bei dem die Medien eine maßgebliche Rolle bei der Urteilsfindung ausübten⁶.

6 Gemäß online-Berichterstattungen der NZZ (*Fontana* 2017; *Gyr* 2017a, 2017b, 2018a, und 2018b) kann festgehalten werden: Nach einem Schuldspruch des

Es bleibt aber festzuhalten, *dass wir ganz einfach zu wenig über Ausmaß und Häufigkeit von Bestätigungsfehlern bei Gerichtsurteilen wissen*. Sofern man mehr darüber erfahren möchte, wäre es notwendig, nicht nur verstärkt Richter in den experimentellen Studien der Psychologie zu berücksichtigen, sondern auch *im Feld* fehlerhafte Urteile zu untersuchen (vgl. Dunkel & Kemme 2016). Erfolgreiche Wiederaufnahmeverfahren wurden in Deutschland letztmals von Peters (1970) untersucht, ohne dass dabei jedoch der Bestätigungsfehler speziell überprüft wurde. Eine vergleichbare, neuere Studie wurde von Gilliéron & Kiliás (2007) in der Schweiz durchgeführt. Für eine Überprüfung gravierenderer Bestätigungsfehler ist es aber nicht nur wichtig, erfolgreiche Wiederaufnahmeverfahren zu analysieren. Um nicht nur die Spitze des Eisbergs zu erfassen, wären zumindest auch die erfolgreichen Berufungsverfahren in die Analyse mit einzubeziehen. Über ein kürzlich in zweiter Instanz getroffenes Urteil am Obergericht Bern soll im Folgenden beispielhaft berichtet werden.

Wegen einseitiger Beweiswürdigung »gekippte« Gerichtsentscheidung: Beschuldigter K wird in erster Instanz (Einzelrichter) zu 9 Monaten Freiheitsstrafe wegen sexueller Nötigung (Art. 189 des Schweizerischen Strafgesetzbuches) verurteilt. Der Anwalt von K geht erfolgreich in Berufung und das zuständige Obergericht (1. Strafkammer des Obergerichts Bern) befindet in seinem Urteil, dass in erster Instanz eine *einseitige Beweiswürdigung* stattgefunden hat.

Was war vorgefallen? Im August 2013 trifft Beschuldigter K zwei Frauen, D und C, am Bahnhof Bern. Der Beschuldigte kennt D bereits und

Urner Obergerichts (erstes Urteil in zweiter Instanz) wurde Ignaz Walker, nachdem das Bundesgericht den Fall ans Obergericht Uri zurückverwiesen hatte, in einem zweiten Urteil vom Vorwurf des Mordversuchs in Mittäterschaft an seiner Ex-Ehefrau freigesprochen. Zu diesem Freispruch kam es nach einem medialen Kesselstreben. Die Presse, insbesondere das Politikmagazin »Rundschau«, hatte sich ohne Wenn und Aber auf die Seite von Ignaz Walker geschlagen und zusammen mit dem Zeugen Sindelic (verurteilter Auftragskiller der Ex-Ehefrau von Walker) eine Komplott-Theorie entwickelt, die die Unschuld von Ignaz Walker belegen sollte. Auch dieses zweite Urteil des Urner Obergerichts wurde ans Bundesgericht weitergezogen, welches das Urteil, verbunden mit einer deutlichen Medienschelte, erneut ans Obergericht zurückverwiesen hat. Im Januar 2018 wird Ignaz Walker vom Urner Obergericht erneut - so wie beim ersten Urteil - schuldig gesprochen, die Ermordung seiner Ex-Ehefrau in Auftrag gegeben zu haben.

in der Folge unternehmen alle drei einen kleinen Ausflug in Bern. Zurück am Bahnhof Bern verabschiedet sich D und der Beschuldigte bietet C an, sie in seinem Auto (Honda Civic) mit in Richtung Fribourg (CH) zu nehmen. Unterwegs halten sie zunächst an einem Stadtfriedhof an, wobei die Gründe hierfür und das, was sich in der Folge abspielte, umstritten sind. Unbestritten ist hingegen, dass C wieder in das Auto einstieg, K mit C erneut zum Bahnhof Bern fuhr und sich die beiden dann dort trennten. Dem Gericht liegen neben diversen subjektiven Beweismitteln auch objektive Beweismittel vor, wie eine Arztbescheinigung, ein Psychotherapeutenbericht, sowie eine Internetdokumentation über den Typ Honda Civic, den der Beschuldigte K fuhr.

Das Obergericht Bern schliesst sich der Würdigung der Vorinstanz in mehreren Punkten *nicht* an, da einerseits die Aussagen des Beschuldigten *nicht generell unglaublich* seien, wie dies vom erstinstanzlichen Gericht beurteilt wurde. So kann es sich beispielsweise, wie von dem Beschuldigten behauptet, bei dem Streitgespräch zwischen C und dem Beschuldigten K durchaus um Probleme der Aufenthaltsbewilligung von C gehandelt haben, da C zum fraglichen Zeitpunkt nur über einen N-Ausweis verfügte. Bei diesem Ausweis handelt es sich nicht um eine reguläre Aufenthaltsbewilligung, wie wohl fälschlich von der ersten Instanz angenommen, sondern lediglich um einen Ausweis, der nur während des laufenden Asylantragverfahrens Geltung hat. Andererseits, so das Obergericht, seien die Widersprüche in den Aussagen von C mit den objektiven Beweismitteln *nur unzureichend* gewürdigt worden.

Diesen Beanstandungen durch das Obergericht sind wir durch eine eingehende Analyse der Strafakten der ersten Instanz nachgegangen⁷. Die vielen Widersprüche zwischen den Aussagen von C (das Auto war verriegelt, so dass ich nicht aussteigen konnte, K klappte beide Sitze nach hinten etc.) mit den technischen Gegebenheiten des fraglichen Autotyps (es gibt weder eine Zentralverriegelung noch ein »Verriegelungsstäbchen« auf der Beifahrerseite oder einen Knopf, mit dem man beide Rückenlehnen nach

7 Die gesamten Strafakten der ersten und zweiten Instanz (Dossiernummer SR 2016 11) wurden uns dankenswerter Weise vom Obergericht Bern am 9.2.2018 vor Ort zur Einsicht zur Verfügung gestellt. Der Entscheid des Obergerichts ist abrufbar unter <http://www.zsg-entscheide.apps.be.ch/tribunapublikation/> unter der o.e. Dossiernummer.

hinten klappen kann etc.) werden von der Erstinstanz durchaus gesehen. So wird angemerkt, dass die Erinnerung von C zu den Funktionen im Auto Fragezeichen aufwerfen würden. In der Gesamtwürdigung wird aber lediglich angemerkt, dass die Frage offenbleiben muss, ob die Türen tatsächlich verschlossen waren und dass C möglicherweise »einfach den falschen Griff beim Öffnen der Tür zu drücken versucht habe«. Dabei wird nicht weiter berücksichtigt, dass C im Verlauf des Geschehens bereits mehrere Male die Autotür geöffnet hat und demnach den Hebel fürs Türöffnen kennen musste. Von der ersten Instanz wird ebenfalls anerkannt, dass die Platzverhältnisse im Honda Civic sehr beschränkt sind (Steuerrad, Mittelkonsole etc.) und es keinen Knopf gibt, mit dem K beide Rückenlehnen hätte nach hinten kippen können. Gleichwohl wird geschlossen, dass der Fahrer problemlos den Hebel an der Aussenseite des Beifahrersitzes hätte bedienen können. Des Weiteren werden Widersprüche zwischen dem Arztbericht und den Aussagen von C zwar berichtet, aber nicht weiter verwertet. So sagt der Arztbericht recht klar, dass bei C keine Fraktur des rechten Zeigefingers vorliegt, wie von C behauptet, insgesamt keine objektivierbaren Prellmarken oder Verletzungsfolgen zu bescheinigen sind und lediglich eine kurzfristige Schonung anzuraten sei. In der Akte der ersten Instanz wird in Bezug auf die von C behaupteten Prellungen und den Aussagen des Arztberichts nur angemerkt, dass für *ungeübte* Augen eine Hautunterblutung bei schwarzer Hautfarbe (C kommt aus Westafrika) nur schlecht zu erkennen sei. Warum C dem Arzt gegenüber zwar sagt, von einem Mann verprügelt worden zu sein, aber nichts von einer sexuellen Belästigung erwähnt, wird in der Beweiswürdigung ebenfalls nicht weiter berücksichtigt.

Natürlich kann an dieser Stelle nicht auf alle Einzelheiten der erstinstanzlichen Beweiswürdigung eingegangen werden. Die obigen Ausführungen sollten aber dennoch eine »unzureichende Beweiswürdigung« hinreichend illustrieren, so wie vom Berner Obergericht beurteilt. Betrachten wir diese unter dem Blickwinkel des Bestätigungsfehlers, so wurden vom erstinstanzlichen Richter (a) *belastende Indizien* zu stark betont (z.B. einschlägige Vorstrafe von K prominent hervorgehoben, vollumfängliches Abstreiten der Vorwürfe als Nachteil ausgelegt) und insbesondere (b) *entlastende Aussagen* verkannt oder abgewertet (z.B. Möglichkeit, dass es beim Vorfall um einen Streit um C's Aufenthaltsbewilligung ging, zu wohlwollende und unkritische Analyse von C's teilweise unlogischen und objektiven Beweismitteln widersprechenden Behauptungen). Ob der Richter oder die Richterin vorzeitig davon überzeugt war, dass K den Straftat-

bestand der sexuellen Nötigung begangen hat, und es deshalb bei der Gewichtung und Analyse des Beweismaterials zu einem Bestätigungsfehler kam, ist natürlich aufgrund der Urteilsbegründung und auch aufgrund unserer Aktenanalyse nicht eindeutig zu erkennen. Es kann daher nur darüber spekuliert werden, dass es zu solch einer frühzeitigen Überzeugung kam, weil K einschlägig vorbestraft war und/oder C als dunkelhäutige Ausländerin einem »Opferstereotyp« entspricht.

Abschließend zu dem Fallbeispiel ist anzumerken, dass das Obergericht Bern in seinem Urteil zu einem Freispruch gelangte. Mit dem Freispruch, so das Obergericht, wird aber nicht der Überzeugung Ausdruck verliehen, dass K's Version des fraglichen Geschehens der tatsächlichen Wahrheit entspricht. Das Gericht erachtet aber die Beweislage als zu wenig klar, um einen Schuldspruch zu rechtfertigen und hat daher in *dubio pro reo* entschieden.

II. Rückschaufehler

Neigen Personen möglicherweise dazu, im Nachhinein ein Ereignis als vorhersehbarer einzuschätzen als im Vorhinein? Diese Frage ist klar zu bejahen. *Fischhoff* (1982) konnte erstmals zeigen, dass Personen ihre Fähigkeit überschätzen, ein eingetretenes Ereignis vorhersehen zu können, und dass sie glauben, auch andere sollten das Ereignis besser vorhersehen können, als es tatsächlich möglich war. Er nannte diesen Urteilsfehler »*hindsight bias*« (Rückschaufehler).

Für Gerichtsentscheidungen ist der Rückschaufehler besonders relevant, da in der Regel geurteilt wird, *nachdem* ein Ereignis X eingetreten ist, jedoch die *ex ante* gegebene Vorhersehbarkeit von X aufgrund gegebener Umstände beurteilt werden soll – unabhängig vom Ausgang der Handlung.

Giroux, Coburn, Harley, Connolly & Bernstein (2016) geben einen Überblick über die verschiedenen Entscheidungen im Rahmen von Gerichtsverfahren, in denen der Rückschaufehler eine Rolle spielen könnte (vgl. Tabelle 1). Über die Fahrlässigkeit wird sowohl im Zivil-, wie im Strafrecht geurteilt, auch wenn ein unterschiedlicher Fahrlässigkeitsbegriff zugrunde liegt. Während im Strafrecht vor allem die subjektive Fahrlässigkeit betrachtet wird (z.B.: Konnte der Wagenlenker eines bewaffneten Banküberfalls zum Zeitpunkt der Tat wissen/vorhersehen, dass eine Person dabei sterben wird?), stellt das Zivilrecht auf die objektive Fahrlässig-

keit ab (z.B.: Konnte ein Durchschnittsbürger anstelle des »[Mit-]Täters« vorhersehen, dass bei einem bewaffneten Banküberfall eine Person sterben wird).

Der Rückschaufehler ist einer der robustesten Urteilsfehler. Ausnahmen konnten nur dann nachgewiesen werden, wenn die Resultate einer Handlung extrem überraschend auftraten, also in keiner Weise damit gerechnet werden konnte (vgl. Müller & Stahlberg 2007; Pezzo 2003). Der Fehler ist dabei umso deutlicher ausgeprägt, je schwerwiegender der Ausgang der betreffenden Handlung ist. Im Zivil- wie im Strafrecht wären dies vor allem Handlungen, die zu schweren Opferschäden geführt haben. Dass der Rückschaufehler mit der Schadenshöhe der Handlung zunimmt, könnte mit einer weiteren Urteilsverzerrung zusammenhängen, dem »Outcome Bias« (vgl. Gino, Moore & Bazerman 2009). Hiernach wird der handelnden Person bei einem negativen (im Vergleich zu einem positiven) Handlungsausgang selbst dann mehr Inkompetenz und moralische Verwerflichkeit zugeschrieben, wenn der jeweilige Ausgang der Handlung rein zufällig zustande gekommen ist.

	Relevante Bereiche	Prognose-Beispiele
Zivilrecht	Medizinische Kunstfehler	Hätte man das Malignom in der Röntgenaufnahme erkennen können?
	Patentrecht	Handelt es sich bei dem Produkt K wirklich um eine neue Erfindung?
Zivil- und Strafrecht	Forensische Untersuchungen/Gutachten	Hat der Polizist den Kläger auf unbefugte Weise mit seiner Waffe verletzt?
	Fahrlässigkeitsbeurteilungen	Hätte man eine Überflutung/einen Dammbbruch mit den verheerenden Folgen vorhersehen können/müssen?
Strafrecht	Strafverfahren	War »mens rea« zum Zeitpunkt der Tat gegeben?

Tabelle 1: Übersicht über verschiedene Bereiche bei Gericht, in denen der Rückschaufehler relevant sein kann. Die Aufstellung folgt mit gewissen Anpassungen der Darstellung von Giroux et al. (2016).

Auf die Rolle des »Outcome Bias« kommen wir nochmals zurück. Betrachten wir zunächst die Frage, ob nur Laien in ihren Urteilen einen Rückschaufehler aufweisen, oder ob er auch für die Gruppe der Richter

nachgewiesen werden konnte. Richterbezogene Untersuchungen zum Rückschaufehler gibt es bisher jedoch nur wenige. Zwei der Untersuchungen wurden von *Guthrie et al.* (2001) und *Rachlinski, Guthrie & Wistrich* (2011) durchgeführt, wobei die Autoren nur in ihrer ersten Studie (*Guthrie et al.* 2001) einen deutlichen Rückschaufehler nachweisen konnten.

In der Studie von *Guthrie et al.* (2001) wurden Richter der USA (magistrate court judges) nach dem Zufallsprinzip drei Gruppen zugeteilt, die man dann jeweils über einen anderen von drei möglichen Ausgängen eines Berufungsverfahrens informierte. Anschliessend sollten alle Richter vorhersagen, für wie wahrscheinlich sie die drei möglichen Ausgänge gehalten hätten, *wenn* ihnen die vorher gegebene Information über den Ausgang des Berufungsverfahrens *nicht* bekannt gewesen wäre. Wurde beispielsweise einer Gruppe gesagt, dass das Urteil bestätigt wurde, so sagten 81,5 % der Befragten, dass sie diesen Ausgang auch am wahrscheinlichsten gehalten hätten. Hingegen sagten nur 27,8 % der Richter, genau diesen Ausgang (Bestätigung des Urteils) am wahrscheinlichsten gehalten zu haben, wenn ihnen zuvor mitgeteilt wurde, dass das Urteil aufgehoben wurde.

In der späteren Untersuchung (*Rachlinski et al.* 2011) konnten die Autoren keinen Rückschaufehler mehr nachweisen. Hier stand in drei verschiedenen Szenarien die Frage der Legalität einer Durchsuchung zur Diskussion, wie z.B. die Durchsuchung der Fracht eines LKWs. Die jeweiligen Gruppen *mit* Angaben zum Ausgang der Untersuchung (Entdeckung von belastendem Material) und *ohne* Angaben zum Ausgang der Durchsuchung, schätzten ihre Legalität nicht signifikant unterschiedlich ein.

In einer weiteren Studie mit einem ähnlichen Design wie in der Studie von *Rachlinski et al.* (2011) konnte allerdings erneut ein recht deutlicher Rückschaufehler aufgezeigt werden. *Oeberst und Goeckenjan* (2016) legten insgesamt 84 deutschen Richtern in einer Vignettenstudie einen Fall zur Überprüfung der Fahrlässigkeit vor, der hier gekürzt wiedergegeben werden soll: Verurteilter L kommt aufgrund seines Straftatenprofils und seiner Persönlichkeitsstörung in eine geschlossene psychiatrische Anstalt. Nach einer bestimmten Zeit der Psychotherapie attestieren ihm alle Verantwortlichen eine positive Entwicklung. Des Weiteren hat L aufgrund einer Anzeige eine neue Beziehung zu einer Frau gefunden, die ihn auch regelmässig besucht. Der Chefarzt und der Oberarzt erlauben ihm einen Freigang für drei Stunden, um mit der neuen Partnerin spazieren zu gehen. Nun werden den Untersuchungsteilnehmenden nach dem Zufallsprinzip entweder *keine* Angaben über den Ausgang des Freigangs (N=41) oder

aber *folgende Angaben* über den Ausgang des Freigangs gegeben (N=43): L entkommt und kann nach sechs Monaten aufgegriffen werden. Während dieser sechs Monate werden acht Personen von L ausgeraubt, wobei eine 80jährige Frau stirbt und drei weitere Personen verletzt werden. Während in der Gruppe der Teilnehmenden *ohne* Angaben des Ausgangs drei Richter (12,2 %) die Entscheidung als fahrlässig beurteilten, stuften in der Gruppe *mit* Angaben des Ausgangs 13 Richter (30,2 %) die Entscheidung als fahrlässig ein⁸.

Vor allem in Bezug auf Fahrlässigkeitsurteile mit hohem Schadensausgang scheint der Rückschaufehler ein ernstzunehmendes Problem sowohl für das Zivil-, wie auch das Strafrecht zu sein. Die Katastrophe bei der Duisburger Loveparade im Juli 2010 ist einer dieser brisanten Fälle. Es starben 21 Menschen und mehr als 650 wurden verletzt. Während *Wolfgang Rabe*, der Chef der städtischen Sicherheitsabteilung, noch einen Tag vor der Katastrophe versicherte, dass die Stadt sehr gut vorbereitet sei, hat nach Eintreten der Katastrophe der britische Panikforscher, *Prof. Keith Still*, den Sicherheitsbeamten Fahrlässigkeit in seinem von der Staatsanwaltschaft in Auftrag gegebenen Gutachten attestiert. Nachdem es zunächst einen Nichteröffnungsbeschluss des Landgerichts Duisburg gab, hat auf grossen Druck der Öffentlichkeit hin (350.000 Online-Unterschriften) und der Beschwerde der Staatsanwaltschaft und der Nebenkläger, das Oberlandesgericht Düsseldorf entschieden, dass das Landgericht ein Strafverfahren gegen zehn Personen eröffnen muss (*Deutsche Presseagentur* 2017).

Grundsätzlich ist es natürlich sehr gut nachvollziehbar und verständlich, wenn Betroffene und auch die Öffentlichkeit gerichtlich geklärt haben möchten, ob ein eingetretener hoher Schaden unvermeidbar gewesen ist, oder aber durch Fahrlässigkeit zustande kam. Gefährlich ist jedoch, dass bei der Fahrlässigkeitsentscheidung sowohl der Rückschaufehler als auch der bereits erwähnte »Outcome Bias« einen Einfluss auf das Urteil nehmen können und dies umso mehr, je höher der entstandene Schaden ist. Während es sich beim Rückschaufehler um einen Erinnerungsfehler handelt, bei dem sich die Person darüber täuscht, was sie geglaubt hat, bevor sie den Ausgang der Handlung kannte (*Fischhoff* 1975), ist der »Outcome

8 Vergleiche auch die Studie von *LaBine & LaBine* (1996), die jedoch nicht mit Richtern durchgeführt wurde.

Bias« in Abgrenzung zum Rückschaufehler nicht auf einen Erinnerungsfehler zurück zu führen. Hier wird je nach Handlungsausgang, auch wenn dieser zufällig oder aufgrund unvorhersehbarer Umstände zustande kam, die handelnden Personen selbst in einem anderen (moralischen) Licht gesehen (vgl. *Gino et al.* 2009).

Aber nicht nur die Allgemeinbevölkerung unterliegt dem »Outcome Bias« (*Oswald, Orth, Aeberhard & Schneider* 2005) und dem Rückschaufehler. Auch in Bezug auf die Rechtsprechung besteht nach *Schweizer* (2005) die Frage, ob ein Nachweis der Einhaltung der ex ante formulierten Sicherheitsnormen (z.B. in Bezug auf die Duisburger Loveparade) überhaupt zur Exkulpierung hinreichend ist, und ob für den Fall fehlender ex ante Normen der »adäquate Kausalzusammenhang« eher ex post beurteilt wird, und es daher auf die ex ante Vorhersehbarkeit kaum noch ankommt. In diesem Sinn würde man zumindest die mit dem Rückschaufehler verbundenen Probleme auf elegante, wenn auch nicht unbedingt gerechte Weise »aushebeln«.

III. Ankereffekt

Angenommen, Sie würden danach gefragt, wie hoch der Eiffelturm sei, und Sie hätten zufälligerweise kurz vorher etwas über den Maintower in Frankfurt/M. (200 m hoch) oder alternativ etwas über das 830 m hohe Hochhaus Burj Khalifa in Dubai gelesen/gesehen, so würde sich die jeweilige Information sehr wahrscheinlich auf Ihre Schätzung auswirken⁹. Dies zumindest dann, wenn Sie nicht in der Lage sein sollten, die annähernd exakte Höhe des Eiffelturms (324 m) aus dem Gedächtnis abzurufen. Würden Sie also im ersten Fall den Eiffelturm als deutlich weniger hoch einschätzen als im zweiten Fall, so wäre dies die Folge eines *Ankereffekts*. Nach *Tversky & Kahneman* (1974, S. 1128) handelt es sich beim Ankereffekt um eine Assimilation/Verzerrung einer Größenschätzung hinsichtlich eines zuvor betrachteten Anfangswerts, bzw. Ankers.

9 In der ursprünglichen Aufgabe von *Tversky & Kahneman* (1974) sollten die Versuchsteilnehmenden den Prozentsatz afrikanischer Mitgliedstaaten in den Vereinten Nationen schätzen.

Für Strafrichter und Strafrichterinnen stellt sich in diesem Zusammenhang die interessante Frage, ob die Höhe ihrer Strafzumessungsurteile möglicherweise auch durch externe, eventuell sogar ganz zufällig vorliegende Vorgaben beeinflusst werden kann. Natürlich gibt es die Strafanträge der Staatsanwaltschaft und auch die der Strafverteidigung, aber hier wäre ein möglicher Einfluss auf das Richterurteil nicht willkürlich, da die Strafanträge in der Regel sehr informativ und aus der Sicht des jeweiligen Antragstellers mehr oder weniger gut begründet sind.

Empirische Befunde zum Ankereffekt zeigen jedoch, dass auch vollkommen irrelevante Urteilsanker einen Einfluss auf die Größenschätzung ausüben (Mussweiler & Strack 2001). In diesem Sinne postulierten Englich, Mussweiler & Strack (2006) in ihrem Aufsatz, dass auch die Strafzumessung der Richter beeinflusst werden kann, wenn ein Strafantrag mehr oder weniger zufällig generiert wurde. Die Autoren führen insgesamt drei Vignettenexperimente durch. In jedem der drei Experimente erhielten die teilnehmenden Richter und/oder Staatsanwälte zunächst eine auf ca. vier Seiten beschriebene Straftat (Vignette), wie z.B. ein Ladendiebstahl oder ein Vergewaltigungsdelikt. Nach dem Zufallsprinzip erhielten die Teilnehmenden dann entweder einen hohen oder niedrigen Urteilsanker. Beim Ladendiebstahlsdelikt entsprach der hohe Urteilsanker 9 Monate unbedingter Freiheitstrafe und der niedrige Urteilsanker 3 Monate bedingter Freiheitstrafe. Im ersten Experiment wurden 42 Richter und Staatsanwälte in jeder Gruppe (hoher vs. niedriger Anker) informiert, dass das Strafmass durch *Journalisten* festgelegt worden war. Im zweiten Experiment wurden 39 Richter und Staatsanwälte darüber informiert, dass die Staatsanwaltschaft das jeweilige Strafmass zum Zweck der Untersuchung *rein zufällig ausgewählt* hat. Im dritten Experiment schliesslich konnten 52 Richter das jeweilige Strafmass selbst durch das *Würfeln zweier Würfel* festlegen. In diesem Fall war das Würfelpaar manipuliert. Die Würfel zeigten entweder immer die Zahlen 3 und 6 (bei hohem Anker) oder aber 1 und 2 (bei niedrigem Anker). In allen drei Experimenten konnte ein klarer Ankereffekt nachgewiesen werden. Nehmen wir den Ladendiebstahl als Beispiel, so lag die abgegebene Strafzumessung bei einem hohen Anker im Durchschnitt drei Monate höher als beim niedrigen Anker.

Bis vor einiger Zeit gingen Forscher aufgrund ihrer experimentellen Ergebnisse davon aus, dass über die verschiedenen Urteilsbereiche hinweg irrelevante Anker einen vergleichbar großen Verzerrungseffekt auf die nachfolgende Größenschätzung ausüben wie relevante Anker (vgl. Muss-

weiler & Strack 2001). Auch English et al. (2006) nahmen in Bezug auf die richterliche Strafzumessung noch an, dass irrelevante Anker (z.B. willkürlich zustande gekommene Strafmasse) einen ähnlich hohen Einfluss ausüben wie relevante Anker (z.B. Strafzumessungsanträge der Staatsanwaltschaft). Im zweiten Experiment einer neueren Studie von Glöckner & English (2015), in der der Einfluss relevanter und irrelevanter Anker direkt miteinander verglichen wurde, konnte jedoch gezeigt werden, dass vor allem bei einem hohen Ankerwert die richterliche Strafzumessung deutlich stärker durch relevante als durch irrelevante Anker beeinflusst wird.

Abschließend möchten wir noch auf eine interessante französische Feldstudie von Geyres, Hilton & Py (2013) hinweisen. Die Autoren untersuchten insgesamt 184 rechtskräftige Strafzumessungsurteile, die in dem Zeitraum von Oktober 2005 bis Mai 2007 von französischen Gerichten in der Gegend um Paris und in Lyon gefällt wurden (ca. 60 pro Jahr). Es handelt sich dabei ausschließlich um Straftaten gegen die öffentliche Ordnung, die mit einer Höchststrafe von 12 Monaten Freiheitsstrafe geahndet werden. Rückfalltäter wurden aus der Analyse ausgeschlossen. Die 184 Urteile waren von insgesamt 12 Richtern an vier verschiedenen Gerichten getroffen worden. Die Strafrichter verfügten über eine unterschiedliche Berufserfahrung, hatten jedoch alle zu Beginn der Untersuchung im Jahr 2005 noch kaum mit Straftaten gegen die öffentliche Ordnung zu tun gehabt.

Der im Jahr 2005 festzustellende recht massive Ankereffekt (64 % der Urteile waren identisch mit dem Strafantrag der Staatsanwaltschaft) nahm im Laufe der Untersuchungszeit ab (2007 waren es noch 43 %). Auffallend an den Ergebnissen war auch, dass sich die Richter innerhalb des Untersuchungszeitraums (2005-2007) recht stark in ihren Strafzumessungsurteilen untereinander angleichen, die Höhe ihrer Urteile also immer weniger streute (vgl. Abbildung 4). Hingegen wiesen die Strafanträge der Staatsanwaltschaft *keine* Veränderung in der Streubreite auf.

Diese Tatsache erlaubte es, eine weitergehende statistische Analyse (ANOVA) durchzuführen.¹⁰ Die Ergebnisse dieser Analyse weisen darauf

10 Es wurde eine Varianzanalyse (ANOVA) mit einem 2 (Ausgeprägtheit des Urteilsankers [moderat, extrem]) x 3 (Strafrichtererfahrung [kurz, mittel, lang]) Design gerechnet. Abhängige Variable war das Ausmaß, mit dem der Richter

hin, dass die Richter umso mehr in ihrem Urteil von den Strafanträgen der Staatsanwaltschaft abwichen, je extremer diese waren und dass vor allem extreme Strafanträge dann mehr ausgeglichen wurden, wenn es sich um berufserfahrenere Richter handelte. Die Autoren schließen aus den Daten, dass vor allem zwei Faktoren den Einfluss des Ankereffekts bei den Richtern verhinderte: Ihre Berufserfahrung und eine im Laufe der Untersuchungsjahre entstandene gerichtsinterne Strafnorm für das fragliche Delikt.

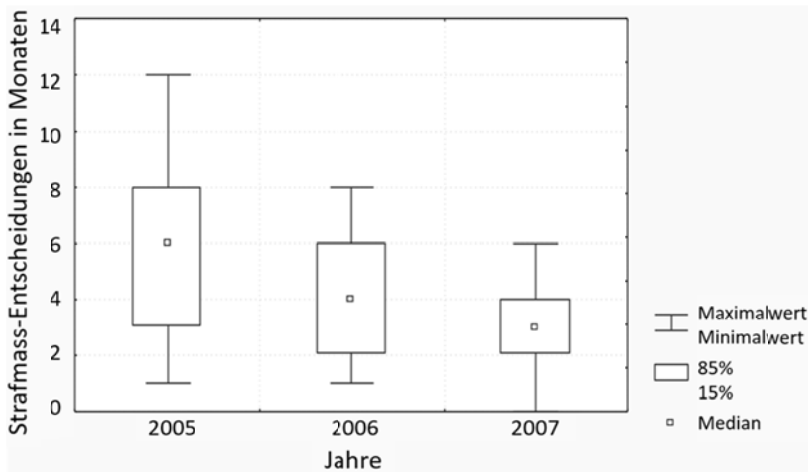


Abbildung 4: Median und Streuung der Strafzumessungsurteile bei Straftaten gegen die öffentliche Ordnung (ca. 60/Jahr) im Verlauf der Untersuchung von Geyres et al. (2013).

Zusammenfassend ist aufgrund der bisher vorliegenden empirischen Befunde zu vermuten, dass externe Urteilsanker das richterliche Strafzumessungsurteil selbst dann beeinflussen können, wenn es sich um irrelevante Anker handelt. Dies könnten beispielsweise Strafforderungen durch die Medien/die Öffentlichkeit sein, oder die Höhe der Strafzumessung von gerade vorausgegangenen Strafurteilen des Richters. Der Ankereffekt könnte

extreme Strafanträge der Staatsanwaltschaft ausglich. Es zeigten sich ein Haupteffekt des Faktors Ausgeprägtheit des Urteilsankers und ein Interaktionseffekt.

dabei stärker sein, wenn Richter insgesamt oder deliktsbezogen noch über wenig Erfahrung als Strafrichter verfügen. Des Weiteren kann vermutet werden, dass Zeitdruck und fehlende oder wenig engagierte Strafverteidigung den Ankereffekt ebenfalls verstärken.

C. Können Urteilsfehler vermindert oder gar vermieden werden?

Aus juristischer Sicht werden immer wieder verschiedene Vorschläge zur Reduktion von Urteilsfehlern gemacht. Hierzu gehört beispielsweise die Forderung, Tonband- oder Videoaufnahmen zum Standard bei Befragungen zu machen, die Hürden für die Wiederaufnahme von Verfahren zu senken, die Verfahrenseröffnung durch ein separates Gremium durchführen zu lassen oder in schwierigen Straffällen einen »advocatus diaboli« hinzuzuziehen. Auf diese Vorschläge werden wir aber nicht weiter eingehen, sondern uns auf Möglichkeiten der Fehlerreduktion aus psychologischer Sicht beschränken.

Da die diskutierten Urteilsfehler im Allgemeinen von den betreffenden Richtern nicht beabsichtigt sind, sondern mehr oder weniger automatisch und unbeabsichtigt passieren, ist es vielleicht verständlich, dass Studien, die den Effekt von *Warnungen* untersuchen, recht einheitlich deren Ineffektivität belegen (*Epley & Gilovich 2005; Wilson, Houston, Etling & Brekke 1996*). Interessant ist jedoch, dass es für die hier diskutierten Urteilsfehler gleichermaßen ein »Gegenmittel« zu geben scheint, welches den Fehler vielleicht nicht gänzlich zum Verschwinden bringt, aber immerhin reduziert. Es handelt sich dabei um ein aktives »consider-the-opposite«. Was bedeutet das für die einzelnen Urteilsfehler?

Der Rückschaufehler kann nachweislich reduziert werden, wenn man die Gründe dafür auflistet, die für einen alternativen Handlungsausgang sprechen und ernsthaft überlegt, welche anderen Ausgänge genauso gut oder sogar wahrscheinlicher gewesen wären (*Slovic & Fischhoff 1977*). Auch der Ankereffekt kann reduziert werden, wenn aktiv Argumente gesucht werden, die gegen die Höhe des extern vorgegebenen Ankers sprechen, oder Gründe dafür gesucht werden, dass der Anker möglicherweise unpassend ist (*Mussweiler, Strack & Pfeiffer 2000; Sanna, Schwarz & Stocker 2002*). Der *Bestätigungsfehler* schliesslich würde am meisten von einem »consider-the-opposite« profitieren, da die Berücksichtigung einer oder mehrerer Alternativen zur Erklärung der Straftat den Fehler sogar beseitigen kann (*Gadenne & Oswald 1986*).

»Sich nicht (frühzeitig) auf einen Tatverdächtigen festzulegen und so lange als möglich nach Alternativerklärungen zu suchen« muss daher oberste Maxime in der Strafverfolgung sein! Diese Forderung klingt dennoch reichlich naiv im Kontext der Strafverfolgung, denn Bestätigungsfehler werden ja *gerade dann* begangen, wenn nicht nur eine beliebige Annahme zu testen ist (vgl. *Gadenne & Oswald 1986*), sondern zugleich eine starke *Überzeugung* in Bezug auf die Schuld besteht, vielleicht zusätzlich Zeitknappheit vorliegt und eine effektive Strafverteidigung fehlt oder »schläft«. Hier wäre also nicht nur eine kognitive Verzerrung zu vermeiden, sondern gegen die Motivation des Richters anzukämpfen, seine bisherige Annahme/Überzeugung zu verteidigen.

Grundsätzlich ist zwar das gesamte Gerichtsverfahren auf dem Prinzip des kontradiktorischen Verfahrens aufgebaut, in dem Richterkollegen, Staatsanwaltschaft und Verteidigung aufgerufen sind, Gegenargumente zu generieren, um eine frühzeitige Festlegung zu verhindern. Aber gerade bei einseitig durchgeführten Ermittlungsverfahren im Vorfeld der Gerichtsverhandlung, bei Einzelrichterentscheidungen, dem mittlerweile hohen Prozentsatz an Strafbefehlen (vgl. *Gilliéron & Kiliás 2007*) oder einem starken medialen Einfluss kann das »consider-the-opposite« speziell in Bezug auf den Bestätigungsfehler auf der Strecke bleiben. Wie oft dies in der Gerichtspraxis aber tatsächlich der Fall ist, wäre durch weitere Forschung dringend zu klären.

Da sich das »consider-the-opposite« als probates Mittel anbietet, die diskutierten Urteilsfehler zu reduzieren, oder gar zu vermeiden, wäre ein systematisches Training dieser Strategie im Rahmen der Weiterbildung aber sicher zu überlegen. Dies ist deshalb besonders wichtig, da ohne ein Training der bloße Vorsatz zum »consider-the-opposite« nicht zuverlässig im Verhalten umgesetzt wird. Erste Programme eines entsprechenden Trainings speziell zur Reduktion des Ankereffekts wurden von *Adame (2016)* erfolgreich getestet.

Literaturverzeichnis

- Adame, B.J.* (2016). Training in the mitigation of anchoring bias: A test of the consider-the-opposite strategy. *Learning and Motivation*, 53(1), 36–48.
- Ask, K. & Granhag, P.A.* (2005). Motivational sources of confirmation bias in criminal investigations: The need for cognitive closure. *Journal of Investigative Psychology and Offender Profiling*, 2(1), 43–63.
- Bahník, Š. & Strack, F.* (2016). Overlap of accessible information undermines the anchoring effect. *Judgment and Decision Making*, 11(1), 92–98.
- Deutsche Presseagentur* (24. März 2017). Love-Parade-Prozess: Darum geht es vor Gericht. Spiegel Online. Abgerufen von <http://www.spiegel.de/panorama/justiz/love-parade-prozess-darum-geht-es-nun-vor-gericht-a-1144606.html>.
- Devine, P.G., Hirt, E.R. & Gehrke, E.M.* (1990). Diagnostic and confirmation strategies in trait hypothesis testing. *Journal of Personality and Social Psychology*, 58(6), 952–963.
- Drizin, S.A. & Leo, R.A.* (2004). The problem of false confessions in the post DNA world. *North Carolina Law Review*, 82(3), 891–1007.
- Dror, I.E., Charlton, D. & Péron, A.E.* (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International*, 156(1), 74–78.
- Dunkel, B. & Kemme, S.* (2016). Fehlurteile in Deutschland: Eine Bilanz der empirischen Forschung seit fünf Jahrzehnten. *Neue Kriminalpolitik*, 28(2), 138–154.
- Englich, B., Mussweiler, T. & Strack, F.* (2006). Playing dice with criminal sentences: The influence of irrelevant anchors on experts' judicial decision making. *Personality and Social Psychology Bulletin*, 32(2), 188–200.
- Epley, N. & Gilovich, T.* (2005). When effortful thinking influences judgmental anchoring: Differential effects of forewarning and incentives on self-generated and externally provided anchors. *Behavioral Decision Making*, 18(3), 199–212.
- Fischhoff, B.* (1975). Hindsight \neq foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1(3), 288–299.
- Fischhoff, B.* (1982). For those condemned to study the past: Heuristics and biases in hindsight. In D. Kahneman, P. Slovic, & A. Tversky (Hrsg.), *Judgment under uncertainty: Heuristics and biases* (S. 335–351). Thousand Oaks, CA: Sage.
- Fontana, K.* (28. April 2017). Falscher Freispruch im Kriminalfall Walker. NZZ. Abgerufen von <https://www.nzz.ch/schweiz/aktuelle-themen/bundesgericht-falscher-freispruch-im-kriminalfall-walker-ld.1289051>.
- Gadenne, V. & Oswald, M.* (1986). Entstehung und Veränderung von Bestätigungstendenzen beim Testen von Hypothesen. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 33(3), 360–374.

- Geyres, B., Hilton, D. & Py, J. (2013). Normalization in the courtroom: Does the formation of a group norm affect how judges adjust prosecutors' demands? *European Journal of Social Psychology*, 43(7), 600–605.
- Gigerenzer, G. & Engel, C. (Hrsg.) (2006). *Heuristics and the law*. Cambridge, MA: MIT Press.
- Gilliéron, G. & Killias, M. (2007). Strafbefehl und Justizirrtum: Franz Riklin hatte Recht! In M. A. Niggli, J. Hurtado Pozo, & N. Queloz (Hrsg.), *Festschrift für Franz Riklin* (S. 379–398). Zürich, Schweiz: Schulthess.
- Gilovich, T., Vallone, R. & Tversky, A. (1985). The hot hand in basketball: On the misperception of random sequences. *Cognitive Psychology*, 17(3), 295–314.
- Gino, F., Moore, D.A. & Bazerman, M.H. (2009). No harm, no foul: The outcome bias in ethical judgments. Harvard Business School NOM Working Paper, Nr. 08-080. Abgerufen von <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.451.913&rep=rep1&type=pdf>.
- Giroux, M.E., Coburn, P.I., Harley, E.M., Connolly, D.A. & Bernstein, D.M. (2016). Hindsight bias and law. *Zeitschrift für Psychologie*, 224(3), 190–203.
- Glöckner, A. & Englich, B. (2015). When relevance matters. Anchoring effects can be larger for relevant than for irrelevant anchors. *Social Psychology*, 46(1), 4–12.
- Greenspan, R. & Scurich, N. (2016). The interdependence of perceived confession voluntariness and case evidence. *Law and Human Behavior*, 40(6), 650–659.
- Guthrie, C.P., Rachlinski, J.J. & Wistrich, A.J. (2001). Inside the judicial mind. *Cornell Law Review*, 86(4), 777–830. Abgerufen von <https://scholarship.law.cornell.edu/cgi/viewcontent.cgi?referer=https://www.google.ch/&httpsredir=1&article=1734&context=facpub>.
- Gyr, M. (30. Oktober 2017a). Ignaz Walkers furioses Schlusswort. NZZ. Abgerufen von <https://www.nzz.ch/panorama/ungluecksfaelle-und-verbrehen/komplotttheorie-endgueltig-beerdigt-ld.2808>.
- Gyr, M. (6. November 2017b). Der Fall Ignaz Walker und die Medien. NZZ. Abgerufen von <https://www.nzz.ch/feuilleton/medien/der-fall-ignaz-walker-und-die-medien-1.18642427?reduced=true>.
- Gyr, M. (22. Januar 2018a). Ignaz Walter schuldig gesprochen. NZZ. Abgerufen von <https://www.nzz.ch/schweiz/ignaz-walker-schuldig-gesprochen-ld.1350121?reduced=true>.
- Gyr, M. (23. Januar 2018b). Der Urner Staatsanwalt im Auge des Medien-Orkans. NZZ. Abgerufen von <https://www.nzz.ch/schweiz/der-unbeirr-bare-ld.1346172?reduced=true>.
- Janisch, W. (17. Mai 2015). Ohne jeden Zweifel: Fehlurteile in Deutschland. *Süddeutsche Zeitung*. Abgerufen von <http://www.sueddeutsche.de/politik/fehlurteile-ohne-jeden-zweifel-1.2479505>.
- Jones, E.E. & Harris, V.A. (1967). The attribution of attitudes. *Journal of Experimental Social Psychology*, 3(1), 1–24.

- Kahneman, D. & Tversky, A.* (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- Kassin, S.M., Dror, I.E. & Kukucka, J.* (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, 2(1), 42–52.
- Kassin, S.M., & Sukel, H.* (1997). Coerced confessions and the jury: An experimental test of the «harmless error» rule. *Law and Human Behavior*, 21(1), 187–203.
- Klayman, J. & Ha, Y.* (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94(2), 211–228.
- Koriat, A., Lichtenstein, S. & Fischhoff, B.* (1980). Reasons for confidence. *Journal of Experimental Psychology: Human Learning and Memory*, 6(2), 107–118.
- Kruglanski, A.W. & Mayseless, O.* (1988). Contextual effects in hypothesis testing: The role of competing alternatives and epistemic motivations. *Social Cognition*, 6, 1–20.
- Kukucka, J. & Kassin, S.M.* (2014). Do confessions taint perceptions of handwriting evidence? An empirical test of the forensic confirmation bias. *Law and Human Behavior*, 38(3), 256–270.
- LaBine, S.J. & LaBine, G.* (1996). Determinations of negligence and the hindsight bias. *Law and Human Behavior*, 20(5), 501–516.
- Langer, E.J.* (1975). The illusion of control. *Journal of Personality and Social Psychology*, 32(2), 311–328.
- Müller, P.A. & Stahlberg, D.* (2007). The role of surprise in hindsight bias: A meta-cognitive model of reduced and reversed hindsight bias. *Social Cognition*, 25(1), 165–184.
- Mussweiler, T. & Strack, F.* (2001). Considering the impossible: Explaining the effects of implausible anchors. *Social Cognition*, 19(2), 145–160.
- Mussweiler, T., Strack, F. & Pfeiffer, T.* (2000). Overcoming the inevitable anchoring effect: Considering the opposite compensates for selective accessibility. *Personality and Social Psychology Bulletin*, 26(9), 1142–1150.
- Narchet, F.M., Meissner, C.A. & Russano, M.B.* (2011). Modeling the influence of investigator bias on the elicitation of true and false confessions. *Law and Human Behavior*, 35(6), 452–465.
- Oeberst, A. & Goekenjan, I.* (2016). When being wise after the event results in injustice: Evidence for hindsight bias in judges' negligence assessments. *Psychology, Public Policy, and Law*, 22(3), 271–279.
- Oswald, M.E.* (2014). Strafrichterliche Urteilsbildung. In T. Bliesener, F. Lösel, & G. Köhnken (Hrsg.), *Lehrbuch der Rechtspsychologie* (S. 244–260). Bern, Schweiz: Verlag Hans Huber.
- Oswald, M.E. & Grosjean, S.* (2004). Confirmation bias. In R. F. Pohl (Hrsg.), *Cognitive illusions: A handbook on fallacies and biases in thinking, judgement and memory* (S. 79–96). Hove, East Sussex: Psychology Press.

- Oswald, M.E., Orth, U., Aeberhard, M. & Schneider, E. (2005). Punitive reactions to completed crimes versus accidentally uncompleted crimes. *Journal of Applied Social Psychology*, 35(4), 718–731.
- Peters, K. (1970). Fehlerquellen im Strafprozess – eine Untersuchung der Wiederaufnahmeverfahren in der Bundesrepublik Deutschland (1. Band – Einführung und Dokumentation). Karlsruhe, Deutschland: Müller.
- Pezzo, M.V. (2003). Surprise, defence, or making sense: What removes hindsight bias? *Memory*, 11(4-5), 421–441.
- Porter, S., ten Brinke, L. & Gustaw, C. (2010). Dangerous decisions: The impact of first impressions of trustworthiness on the evaluation of legal evidence and defendant culpability. *Psychology, Crime & Law*, 16(6), 477–491.
- Pronin, E., Lin, D.Y. & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. *Personality and Social Psychology Bulletin*, 28(3), 369–381.
- Rachlinski, J.J. (2000). Heuristics and biases in the court: Ignorance or adaptation? *Oregon Law Review*, 79(1), 61–102.
- Rachlinski, J.J., Guthrie, C. & Wistrich, A.J. (2011). Probable cause, probability, and hindsight. *Journal of Empirical Legal Studies*, 8(1), 72–98.
- Rassin, E. (2017). Initial evidence for the assimilation hypothesis. *Psychology, Crime & Law*, 23(10), 1010–1020.
- Sanna, L.J., Schwarz, N. & Stocker, S.L. (2002). When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(3), 497–502.
- Schmittat, S.M. & English, B. (2016). If you judge, investigate! Responsibility reduces confirmatory information processing in legal experts. *Psychology, Public Policy, and Law*, 22(4), 386–400.
- Schweizer, M.D. (2005). Kognitive Täuschung vor Gericht (Dissertation Universität Zürich). Zürich, Schweiz: Universität Zürich.
- Slovic, P. & Fischhoff, B. (1977). On the psychology of experimental surprises. *Journal of Experimental Psychology: Human Perception and Performance*, 3(4), 544–551.
- Smalarz, L., Madon, S., Yang, Y., Gyll, M. & Buck, S. (2016). The perfect match: Do criminal stereotypes bias forensic evidence analysis? *Law and Human Behavior*, 40(4), 420–429.
- Snyder, M. & Swann, W.B., Jr. (1978). Hypothesis-testing in social interaction. *Journal of Personality and Social Psychology*, 36(11), 1202–1212.
- Trope, Y. & Bassok, M. (1982). Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology*, 43(1), 22–34.
- Trope, Y. & Liberman, A. (1996). Social hypothesis testing: Cognitive and motivational mechanisms. In E.T. Higgins & A. W. Kruglanski (Hrsg.), *Social psychology. Handbook of basic principles* (S. 239–270). New York, NY: Guilford Press.

- Tversky, A. & Kahneman, D.* (1974). Judgment under uncertainty: Heuristics and biases. *Science* 185(4157), 1124–1131.
- Volbert, R.* (2013). Falsche Geständnisse. Über die möglichen Auswirkungen von Voreinstellung, Vernehmung und Verständigung. *Forensische Psychiatrie, Psychologie, Kriminologie*, 7(4), 230–239.
- Vrij, A., Mann, S., Kristen, S. & Fisher, R.P.* (2007). Cues to deception and ability to detect lies as a function of police interview styles. *Law and Human Behavior*, 31(5), 499–518.
- Wason, P.C.* (1960). On the failure to eliminate hypothesis in a conceptual task. *Quarterly Journal of Experimental Psychology*, 12(3), 129–140.
- Wilson, T.D., Houston, C.E., Etling, K.M. & Brekke, N.* (1996). A new look at anchoring effects: Basic anchoring and its antecedents. *Journal of Experimental Psychology: General*, 125(4), 387–402.
- Zuckerman, M., Knee, C.R., Hodgins, H.S. & Miyake, K.* (1995). Hypothesis confirmation: The joint effect of positive test strategy in acquiescence response set. *Journal of Personality and Social Psychology*, 68(1), 52–60.

