

Zum Stand, zur Notwendigkeit und zu den Aussichten der Prozessfehler- und Fehlurteilsforschung

Von Professor Dr. Ralf Köbel, München,
Professor Dr. Jens Puschke LL.M., Marburg,
Professor Dr. Tobias Singelstein, Bochum

Fehlurteile wurden im bundesdeutschen Wissenschaftsdiskurs lange kaum thematisiert, weder aus einer strafprozessualen noch aus einer kriminalsoziologischen Perspektive. Die vor Jahrzehnten hierzu durchgeführten (und zum Teil durchaus bemerkenswerten) Studien¹ haben keine Nachfolger gefunden, auch als sie aufgrund der juristischen und forensischen Entwicklungen ihre Aktualität zu verlieren begannen. Dies wiederum nährte die unterschwellige Überzeugung, dass eine untersuchungsbedürftige Fehlurteilsproblematik hierzulande kaum bestehe – eine Annahme, die beispielsweise noch 2015 in der schulterzuckenden Gleichgültigkeit sichtbar wurde, mit der die sog. »StPO-Expertenkommission« das geltende Wiederaufnahmerecht für das der Problematik angemessene Rechtsregime hielt.² Zumindest in Ansätzen wird dem Thema inzwischen allerdings doch wieder eine gewisse Aufmerksamkeit zuteil. Zwei wissenschaftliche Schwerpunkthefte³ und drei Tagungen⁴ befassten sich in den letzten fünf Jahren mit fehlerurteilsbezogenen Fragen. Auch liegen inzwischen die Ergebnisse von drei kleinen, dafür aber einschlägigen und vor allem aktuellen Erhebungen vor.⁵ Da das Wissen um die Phänomenologie und Prävalenz von Fehlurteilen in Deutschland aber trotzdem noch ausgesprochen defizitär ist, muss zur Problemabschätzung auf internationale Befunde zurückgegriffen und deren Übertragbarkeit erwogen werden. Folglich präsentiert der vorliegende Beitrag einen Überblick über den Stand der internationalen Forschung, prüft deren Relevanz für den deutschen Strafprozess und gibt einen Ausblick auf die erforderlichen juristischen und empirischen Analysen.

-
- 1 Das betrifft vor allem die Auswertung von über 1.100 Wiederaufnahmeverfahren aus den Jahren 1951–1964, deren Ergebnisse von *Peters* unter dem Titel »Fehlerquellen im Strafprozeß« in drei Bänden (1970, 1972, 1974) veröffentlicht wurden. Hierzu sowie zu einigen weiteren, kleineren Untersuchungen aus dieser Zeit vgl. den Überblick bei *Dunkel/Kemme* NK 28 (2016), 138 (145 f.).
 - 2 Bericht der Expertenkommission zur effektiveren und praxistauglicheren Ausgestaltung des allgemeinen Strafverfahrens und des jugendgerichtlichen Verfahrens, 2015, S. 168 ff. (unter: https://www.bmju.de/SharedDocs/Downloads/DE/PDF/Abschlussbericht_Reform_StPO_Kommission.html – Aufruf v. 4.12.2018). Für eine deutlich andere Tendenz vgl. etwa *Eschelbach/Geipel/Hettinger/Meller/Wille* GA 2018, 238 (241 ff., 245 ff.).
 - 3 Vgl. Heft 4/2013 von *Forensische Psychiatrie, Psychologie, Kriminologie* und Heft 1/2016 von *Recht & Psychiatrie*.
 - 4 Einmal die Sommerveranstaltung des AK-»Psychologie im Strafverfahren«, die unter dem Titel »Fehlerquellen im Strafverfahren – Ursachen von Fehlurteilen« am 11. und 12.8.2017 in Bad Saarow stattfand, zum anderen eine Veranstaltung des Vereins »Deutsche Strafverteidiger« am 30.11. und 1.12.2018 in Berlin sowie eine große Tagung am 23. und 24.11.2017 am Bielefelder ZiF (dazu *Barton/Dubelaar/Köbel/Lindemann* Vom hochgemuten voreiligen Griff nach der Wahrheit, 2018, sowie der Tagungsbericht bei *Brückmann/Nießing/Sommer/van Toor* MschrKrim 2017 [100], 483 ff.).
 - 5 Vgl. *Hoffmann/Leuschner* Rehabilitation und Entschädigung nach Vollstreckung einer Freiheitsstrafe und erfolgreicher Wiederaufnahme, 2017; *Böhme* Das strafgerichtliche Fehlurteil – Systemimmanenz oder vermeidbares Unrecht?, 2018; *Dunkel* Fehlentscheidungen in der Justiz, 2018.

I. Fehltritte als Untersuchungsgegenstand

1. Begriff und Feststellbarkeit

Jede empirisch fundierte Auseinandersetzung mit Fehltritten ist zu Anbeginn mit der Frage nach den Merkmalen und der Feststellbarkeit von Fehltritten konfrontiert. Hierbei geht es um eine *definitorische* Festlegung, die anhand ihrer wissenschaftlichen Zweckmäßigkeit und analytischen Eignung erfolgen sollte. Je nach Forschungsinteresse können ganz unterschiedliche Konstellationen von fehlerhaften Entscheidungen als »Fehlurteil« bezeichnet und problematisiert werden. Da es im hiesigen Zusammenhang (auch) um eine Auseinandersetzung mit der ausländischen Forschung gehen soll, liegt es insofern nahe, zunächst den konventionellen, international verwendeten Sprachgebrauch heranzuziehen (zu weitergehenden Überlegungen unten IV.2.). Dieser Fehltrittsbegriff rekurriert auf Strafurteile mit einer besonders hochgradigen, evidenten Fehlerhaftigkeit.⁶ Fehltritte werden in den vorliegenden Arbeiten nämlich fast durchgehend – meist aber ohne dies zu explizieren – mit solchen Strafurteilen gleichgesetzt, denen erstens ein unzutreffender Entscheidungssachverhalt zugrunde liegt und in denen dies zweitens eine überhöhte oder unbegründete Strafe nach sich zieht.⁷ Dieses Verständnis bezieht sich im Grunde also auf prozessuale Vorgänge, bei denen eine Person wegen eines Verhaltens verurteilt wird, das eigentlich gar nicht stattgefunden hat oder das einer anderen Person zuzurechnen ist.⁸

Eine derart zugeschnittene Fehltrittskategorie entspricht – indem sie die Möglichkeit einer »richtigen« (i. S. von »wahrheitswertigen«) Sachverhaltserhebung voraussetzt – dem korrespondenztheoretischen Wahrheitsverständnis der Strafprozessordnung, die eine Übereinstimmung von prozessual festgestelltem Geschehen und Realereignis anstrebt, auch wenn sie rechtspraktisch bereits eine ohne vernünftigen Zweifel bleibende richterliche Überzeugung ausreichen lässt.⁹ Tatsächlich lässt sich das »echte« Fallgeschehen jedoch schon hinsichtlich seiner kognitiven und emotionalen Sinngebungen gar nicht ergründen und in objektiver Weise rekonstruieren.¹⁰ Alles, was ein Gericht erkenntnistheoretisch erreichen kann, ist eine intersubjektiv akzeptable und akzeptierte Ereignisrekonstruktion.

Mit Blick auf den Fehltrittsbegriff bedeutet dies, dass jeder Akteur, der eine Sachverhaltsfeststellung und eine darauf gründende Entscheidung eines Gerichts als richtig oder falsch beurteilen soll, in gleicher Weise diesen Begrenzungen

6 Zum Folgenden stellvertretend *Huff/Rattner/Sagarin Crime and Delinquency* 32 (1986), 518 (519); *Cassell Ariz. L. Rev.* 60 (2018), 815 (819).

7 Genau genommen ließen sich auch unzutreffende Gefahrenprognosen – etwa eine unbegründete Maßregelordnung nach fälschlicher Bejahung der täterseitigen Gefährlichkeit – als Unterfälle des gängigen Fehltrittskonzepts auffassen. Sie werden jedoch üblicherweise nicht in diesem Zusammenhang, sondern als eigenständige Materie diskutiert (für einen Überblick vgl. *Eisenberg/Kölbel Kriminologie*, 7. Aufl. 2017, § 21 Rn 14 ff.).

8 Zur Unterscheidung zwischen »no crime«- / »real crime, wrong perpetrator«-convictions und zu den hierfür typischen Besonderheiten eingehend *Henry American Criminal Law Rev.* 55 (2018), 665 ff.

9 Etwa BGH NSZ 1988, 236; NSZ 1994, 554.

10 Hinzuweisen ist hier auch auf die Begrenztheit der Beweismethoden, auf die Verfremdungswirkung einer normorientierten Fallbildung, auf die Verzeichnungseffekte der Rechtssprachlichkeit – allesamt unhintergehbare Wahrheitshindernisse. Näher zum Ganzen beispielsweise *Stamp Die Wahrheit im Strafverfahren*, 1998, S. 139 ff.; *Gerson Das Recht auf Beschuldigung*, 2016, S. 683 ff.

unterliegt und ebenfalls keine »wahre« Geschehensversion präsentieren kann.¹¹ Was ihm allerdings – etwa dank höherer Sorgfalt, größerer Zeitressourcen, zusätzlicher Informationsquellen oder besserer Beweisverfahren – bei gegebenem Anlass oftmals möglich sein wird, ist die – mit unterschiedlichem Objektivierungsgrad erfolgende¹² – Falsifizierung einer gerichtlichen Sachverhaltsfeststellung. Ein so widerlegtes Urteil verliert dann seine intersubjektive Akzeptanzfähigkeit und wird (fallkonkret) als falsch markiert. In diesem Sinn sind Fehlerurteile jedenfalls auf einem Plausibilitätsniveau, das in den sozialen und justiziellen Alltagspraxen hinreichende Überzeugungskraft hat, retrospektiv prinzipiell *feststellbar*.¹³

2. Wege der wissenschaftlichen Identifizierung

Eine andere Frage ist, welche Urteile als Fehlerurteile *gelten*. Fehlerurteile weisen sich nicht selbst als fehlerhaft aus, sondern stellen bis zu ihrer Widerlegung im gesellschaftlichen Zusammenhang sozial unauffällige »Normalentscheidungen« dar.¹⁴ Damit ein Urteil als falsch behandelt wird, bedarf es eines Ereignisses, das es als falsch einstuft. Im Alltagsleben und im Rechtssystem wird eine urteilsabweichende, institutionelle Entlastungsentscheidung (»exoneration«) als hierfür maßgeblicher Akt angesehen. Als Fehlerurteile begreift man danach jene richterlichen Entscheidungen, die durch einen anderen justiziellen Akteur nach Rechtskraft kassiert und nach einer abweichenden Sachverhaltsbildung durch einen Freispruch oder einen substanziell gemilderten Strafspruch ersetzt werden. Dies sollte jedoch nicht den Blick darauf verstellen, dass prinzipiell auch andere markierende Ereignisse ausreichend sein können. So wird es vielfach Akteursgruppen geben (etwa die Verteidigung, das Umfeld des Verurteilten, die Presse usw.), die manches Strafurteil aus eigener Sachverhaltseinschätzung und ohne Entlastungsentscheidung für ein Fehlerurteil halten.

In Ansehung dessen bestehen für die (sozial-)wissenschaftliche Untersuchung von Fehlerurteilen verschiedene Möglichkeiten, um Entscheidungen, die als Fehlerurteile gelten, zu *identifizieren* und damit einer Analyse zu unterziehen: Entweder werden gerichtliche Entlastungsentscheidungen zusammengetragen, um sodann die darin korrigierten Erst- bzw. Ausgangsurteile auszuwerten, oder man befasst sich mit jenen Strafurteilen, die in den Augen von Prozessakteuren oder Experten fehlerhaft sind. Für die Wahl zwischen diesen Varianten sind vielfach forschungspragmatische Aspekte bestimmend, wie etwa Feldzugänge, vorhandene Daten und andere Machbarkeitsaspekte. Zu berücksichtigen ist aber auch, welche Aussagen und Interpretationen die verschiedenen Ansätze erlauben – und welchen Begrenzungen sie unterliegen.

¹¹ Dazu *Kotsoglou JZ* 2017, 123 (131 f.).

¹² Bei einem nachträglichen »Unschuldsbeweis«, der auf einem DNA-Abgleich beruht, ist der Objektivierungsgrad wesentlich höher als bei einer neuen entlastenden Zeugenaussage.

¹³ In der Sache ähnliche Argumentation bei *Barton FS Eisenberg*, 2019, S. 15 (17 ff.); abweichend aber *Kotsoglou JZ* 2017, 123 (129, 132): konkretes Fehlerurteil bei rechtswissenschaftlicher Betrachtung nicht feststellbar.

¹⁴ Das bedeutet insbesondere, dass die verurteilte Person gesellschaftlich als schuldige und nicht als unschuldige Person behandelt wird.

3. Bedingungen von Helffeldmethoden

Die meisten Studien wählen den Zugang zu Fehlurteilen über dokumentierte Entlastungsentscheidungen. Ihr Untersuchungsmaterial beruht auf entdeckten und institutionell revidierten Falschentscheidungen, was methodisch als Fehlurteilshelffeld bezeichnet werden kann.¹⁵ Dieser Ansatz hat einen spezifischen Vorzug: Weil eine justizielle Korrektur rechtskräftiger Strafurteile typischerweise erst nach einem aufwändigen Zweitverfahren erfolgt – in Deutschland nach einer hochschwelligten Wiederaufnahme und einer neuen gerichtlichen Anklageprüfung –, weist sie mit relativ hoher *Verlässlichkeit* darauf hin, dass die ursprüngliche Verurteilung tatsächlich nach intersubjektiv geteilten (Praxis-)Maßstäben fehlerhaft war. Die hierüber identifizierten Urteile gelten also nicht nur als Fehlurteile (soeben 2.), sondern *sind* im oben (1.) behandelten Sinne sehr wahrscheinlich auch solche.

Andererseits bestehen hierbei aber auch ganz erhebliche Probleme. Die erste Schwierigkeit ist forschungspraktischer Art und betrifft den *Zugang* zu den Entlastungs- sowie den darin korrigierten Ausgangsentscheidungen. »Exonerations« werden als solche in Registern und Justizstatistiken sehr unzulänglich erfasst, was sowohl eine statistische Auswertung als auch die Ermittlung der konkreten Verfahren (anhand der Aktenzeichen) drastisch erschwert. Dies gilt auch für Deutschland. So enthalten weder das Bundeszentral- noch das Staatsanwaltschaftliche Verfahrensregister die erforderlichen Daten zur Abänderung rechtskräftiger Urteile.¹⁶ Die amtlichen Angaben über Häufigkeit und Ausgang von Berufungs- und Revisionsverfahren¹⁷ sind von vornherein zu unspezifisch und allenfalls für Aussagen zu (rechtsmittelförmig) verhinderten Fehlurteilen aufschlussreich. Soweit behördliche Informationen über die Gewährung von Haftentschädigung zugänglich sind,¹⁸ bleiben hierin sämtliche Fehlurteile, in deren Folge es nicht zu einer vollstreckten Freiheitsstrafe kam, unberücksichtigt;¹⁹ auch wird eine Entschädigung nicht stets beantragt und gewährt. Ähnlich verhält es sich bei erfolgreichen Wiederaufnahmeverfahren, die zu einem Frei- oder milderen Schuldspruch führen. Zwar handelt es sich dabei exakt um die hier maßgeblichen Entlastungsvorgänge, doch werden die dahingehenden Daten amtlich nur bedingt regis-

15 Hierzu und zum Folgenden auch *Kölbel* in: Barton/Dubelaar/Kölbel/Lindemann (Fn 4), S. 34 ff.

16 Im BZR kommt es in solchen Fällen zur Löschung oder Überschreibung der korrigierten Ursprungsentscheidung. Zur Nichtidentifizierbarkeit der fraglichen Verfahren in den Daten der Staatsanwaltschaft vgl. *Hoffmann/Leuschner* (Fn 5), S. 25.

17 Dazu im vorliegenden Zusammenhang *Jehle* *Forensische Psychiatrie, Psychologie, Kriminologie* 7 (2013), 220 (225 ff.).

18 Vgl. etwa <https://de.statista.com/statistik/daten/studie/241461/umfrage/unschuldig-inhaftierte-und-entschaedigte-hafttage-in-deutschland/> (Aufruf v. 4.12.2018).

19 Die mit Abstand meisten Haftentschädigungen werden für unbegründete Untersuchungshaft gezahlt (vgl. etwa BayLT-Drs. 17/2801, S. 2 f.). Für die Zeit von 1971 bis 1994 weist die früher geführte StrEG-Statistik einen Jahresdurchschnitt von bundesweit 13 Entschädigungen nach Strafhaft aus. Für 1997 bis 2014 lagen die Werte für sechs bzw. sieben Bundesländer insgesamt zwischen null bis vier pro Jahr (*Böhme* [Fn 5], S. 65 ff.). Auf der Basis der Entschädigungsdaten konnten *Hoffmann/Leuschner* (Fn 5), S. 22 ff. daher (trotz ergänzender Recherchen bei Anklagebehörden und in Medienberichten) auch nur sehr wenige korrigierte Fehlurteile zusammentragen. Wegen der schwierigen Aktenbeschaffung handelte es sich bundesweit von 1990 bis 2017 um lediglich 29 Verfahren mit Freisprüchen nach (Teil-)Vollstreckung einer Freiheits- oder Jugendstrafe. Der Grad, mit dem damit die tatsächliche »exonerations«-Anzahl erfasst wurde, blieb unklar (*Hoffmann/Leuschner* [Fn 5], S. 27).

triert. Die Geschäftsstatistik weist allein die Häufigkeit der Wiederaufnahmeanträge aus,²⁰ nicht aber deren Ergebnisse. Deshalb kann allenfalls eine Stichprobe der registrierten Wiederaufnahmeverfahren auf die hierin enthaltenen erfolgreichen Anträge durchgesehen werden, um so einen Bestand an Entlastungs- bzw. fehlerhaften Ursprungsentscheidungen der näheren Untersuchung zugänglich zu machen.²¹

Einige der vorliegenden deutschen Arbeiten sind diesen Weg gegangen.²² Allerdings ist die Aussagekraft ihrer Ergebnisse stark limitiert. Abgesehen von der Frage, ob die ausgewerteten Wiederaufnahmeverfahren wenigstens für alle erfolgreichen Wiederaufnahmeanträge repräsentativ sind,²³ besteht bei einer solchen Vorgehensweise ein zweites, ganz *grundlegendes* Problem: Wie bei allen anderen Zugängen über das Fehlerurteilshellfeld können allein tatsächlich aufgehobene Fehlerurteile einbezogen werden – nicht aber jene, bei denen eine Wiederaufnahme und/oder Korrektur entweder an den hochschwierig gehandhabten Voraussetzungen scheitern oder wegen eines Mangels an Interesse, Ressourcen oder Erfolgsaussichten gar nicht erst angestrengt werden. Gerichtliche Entscheidungen, die Fehlerurteile sind, aber nicht offiziell als solche gelten (oben 2.), bleiben damit ausnahmslos im Dunkelfeld und können nicht untersucht werden.²⁴ Da davon ausgegangen werden muss, dass die Zusammensetzung des Fehlerurteilshellfelds durch *Selektionsprozesse* massiv geprägt wird,²⁵ haben derartige Untersuchungen eine sehr spezifische Auswahl an Verfahren zum Gegenstand. Aufgrund dessen lassen sich weder Aussagen zum Dunkelfeld treffen noch belastbare Annahmen zu Größe und Struktur des Problems insgesamt formulieren.

4. Bedingungen von Dunkelfeldmethoden

In Anbetracht dieser methodologischen Problematik werden mittlerweile Verfahren eingesetzt, die zumindest teilweise auch das Fehlerurteilsdunkelfeld erschließen wollen. Diese Arbeiten untersuchen Entscheidungen, die nicht wegen ihrer amtlichen Revidierung, sondern kraft einer nicht-institutionellen Beurteilung als Fehlerurteile gelten (oben 2.). Es werden also beispielsweise Urteile analysiert, die nach Ansicht von Verfahrensbeteiligten (Verurteilte, Verteidiger) fehler-

20 Zur Diskussion solcher Daten im vorliegenden Zusammenhang vgl. für Bayern *Kaspar/Arnemann R&P* 34 (2016), 58 (62); BayLT-Drs. 17/658, S. 4 ff.

21 Aber auch dann führen die Unzulänglichkeiten der Aktenregistrierung und -archivierung erfahrungsgemäß zu erheblichen Problemen bei der Aktenzeichenermittlung und Aktenbeschaffung, wodurch die Vollständigkeit des Untersuchungsmaterials stark beeinträchtigt wird (dazu die Schilderungen bei *Hoffmann/Leuschner* [Fn 5], S. 26; *Dunkel* [Fn 5], S. 170 ff.).

22 *Kiwit* Fehlerurteile im Strafrecht, 1965; *Peters* (Fn 1); *Dunkel* (Fn 5), S. 179 ff.

23 Die Frage ist für die eben genannten Arbeiten (Fn 22) zu verneinen.

24 Wenn Urteile erst durch entsprechende gesellschaftliche Bewertungsakte zu Fehlerurteilen werden (dazu bei Fn 14), ist es streng genommen inkonsequent, bei jenen Urteilen, bei denen eine solche Bewertung gut begründbar wäre, aber aus irgendwelchen Gründen unterbleibt, von einem Fehlerurteilsdunkelfeld zu sprechen. Aber ohne diese Inkonsequenz verlöre man die Möglichkeit, diese Urteile überhaupt zu thematisieren. Hierzu am Parallelproblem des Dunkelfelddeliktes, das es für einen konstruktivistischen Wahrheits- und Deliktsbegriff eigentlich ebenfalls nicht »gibt«, siehe *Eisenberg/Köbel* Kriminologie, § 8 Rn 9, § 16 Rn 9 f.

25 Beispiele: Das Interesse daran, ein Urteil anzugreifen, sollte durch die verhängte Strafart/-höhe beeinflusst sein. Und die Wahrscheinlichkeit, dass die Verurteilten neue Beweise vorlegen können, dürfte ebenso wie die Fehleranfälligkeit der Ausgangsverfahren deliktsabhängig variieren (siehe auch Fn 55).

haft sind.²⁶ Dadurch geraten viele Strafurteile in den Blick, bei denen es (noch) nicht zu einer gerichtsseitigen Entlastung kam. Dass es dazu kommen könnte (bzw. sollte) und dass die jeweilige Tatsachengrundlage unhaltbar ist, stellt hier aber lediglich eine subjektive Einschätzung dar. Untersucht werden auf diese Weise also nur »verdächtige« oder »anrühige« Verurteilungen.²⁷

Die intersubjektive Akzeptabilität der Fehlerurteilseinstufung kann demgegenüber merklich erhöht werden, wenn eine Untersuchung einen Fallbestand nach wissenschaftlichen Standards überprüft und auf Basis der eigenen neutralen Rechtsexpertise (oder der von hinzugezogenen Experten) jene Strafurteile herausfiltert, die in tatsächlicher Hinsicht unzutreffend sind. Dies ist allerdings mit einem außerordentlichen Aufwand verknüpft, da die Häufigkeit von Fehlerurteilen im hier bezeichneten Sinn vermutlich relativ gering ist, so dass eine aussagekräftige Anzahl nur durch Überprüfung einer sehr großen Verfahrensstichprobe ermittelt werden kann. Folglich sind Untersuchungen dieser Art selten.²⁸ Ohnehin lassen sich die jeweiligen Verfahren dabei meist nur auf Aktenbasis beurteilen, weil kein zusätzliches Beweismaterial (neue Aussagen, neue Spurenauswertungen usw.) verfügbar ist.²⁹ Deshalb beziehen sich solche Arbeiten letztlich auf Urteile, die streng genommen *aktenwidrig* sind und sich mit dem Bestand an Fehlerurteilen lediglich überschneiden.

II. Stand der US-amerikanischen Fehlerurteilsforschung

1. Prävalenzbestimmung

Die skizzierte Methodenproblematik ist grundsätzlicher Natur und prinzipiell unabhängig von spezifischen Ausgestaltungen der Rechtsordnungen. Die Fehlerurteilsforschung hat sich insbesondere in den USA gleichwohl als eine eigenständige »area of academic study« konstituiert, die eine bemerkenswerte Anzahl systematischer Auswertungen hervorbringt und sich methodisch zunehmend sozialwissenschaftlicher Mittel bedient.³⁰ Die Arbeitsgrundlage bilden meist die kontinuierlich wachsenden *Falldokumentationen*, die von Bürgerrechtsorganisationen und Wissenschaftsakteuren aufgebaut worden sind.³¹ Neben den Falldatenbanken

26 Für Beispiele siehe Fn 47 f. sowie im Grunde auch *Böhme* (Fn 5), S. 184 ff.

27 Ganz in diesem Sinn auch *Wagenaar/van Koppen/Crombag* Anchored narratives, 1993, S. 11 ff., die jene Verfahren, in denen die Verteidiger trotz der Verurteilung fest von der Unschuld des Mandanten überzeugt waren, als einen Typ »dubioser Fälle« analysieren – wobei echte Falschverurteilungen Unschuldiger davon nur eine Teilmenge darstellten.

28 Dergleichen hat man in Deutschland bislang auch nur für eine spezielle Sonderkonstellation durchgeführt: nämlich in Gestalt von Aktenuntersuchungen, die unter sachverständiger Mitwirkung zeigen, dass und wie oft eine naheliegende Schuldfähigkeit weder von der Staatsanwaltschaft noch dem Gericht erkannt worden ist und die Begutachtung und eine Anwendung der §§ 20, 21 StGB demgemäß unterblieb. Allerdings waren die Stichproben hier jeweils vorselektiert. Vgl. zu den Arbeiten m. w. N. *Eisenberg/Köbel* Kriminologie, § 30 Rn 57; siehe auch Fn 67; zur Problematik aus der internationalen Fehlerurteilsforschung für Norwegen *Stridbeck/Magnussen* Criminal Law Quarterly 58 (2012), 267 (280); näher auch *Grøndahl/Stridback* Criminal Behaviour and Mental Health 26 (2016), 212 ff.

29 Dies ist nur in besonderen Konstellationen, in denen die Experteneinschätzung sogar weitgehend objektiviert werden kann, anders (zu einem seltenen Beispiel Fn 39).

30 So bereits die Bilanz von *Leo J. Contemporary Criminal Justice* 21 (2005), 201 (204), demzufolge bereits die Entwicklung einer »criminology of wrongful conviction« anstehe.

31 Einschlägige amtliche Zahlen sind in den USA noch weniger als in Deutschland vorhanden.

des »Death Penalty Information Center«³² und des »Innocence Projects«³³ betrifft dies insbesondere ein umfangreiches, aus verschiedenen Quellen (anwaltliche Meldungen, Presseberichte usw.) gespeistes landesweites Verzeichnis mit inzwischen über 2.300 erfassten Entlastungsentscheidungen (»National Registry of Exonerations«).³⁴ Durch die Prominenz dieser Datenbanken und die Vernetzung der Trägerorganisationen ist sichergestellt, dass stattfindende »exonerations« darin mit hoher Wahrscheinlichkeit registriert werden.

Allerdings besteht auch hier das Problem eines unbestimmbaren Bereichs nicht entdeckter und nicht revidierter Fehlerurteile. Aussagen zu deren »wirklicher« Verbreitung lassen diese Fallsammlungen also (trotz ihres Umfangs) keineswegs zu (oben I.3.). Deshalb konzentrieren sich die bislang erfolgten Prävalenzberechnungen auf jenen Teilbereich der Strafjustiz, bei dem anzunehmen ist, dass stattfindende Fehlerurteile mit substanzieller Häufigkeit aufgedeckt und korrigiert werden und deshalb zu hohen Anteilen in die Fallsammlungen eingehen. Diese Sonderbedingungen sind namentlich bei Verurteilungen zur *Todesstrafe* gegeben, weil deren Vernichtungskarakter zu einem gesteigerten Überprüfungsinteresse führt. Außerdem ist im Fall von rechtskräftigen, aber als problematisch interpretierten Urteilen der Mobilisierungsgrad von Anwälten, diversen Organisationen, aber auch der Justiz oft so stark, dass in diesem Teilbereich der strafgerichtlichen Praxis von einem vergleichsweise geringen Fehlerurteilsdunkelfeld ausgegangen wird.³⁵

Auf der Basis dieser Prämissen ergab sich aus der Anzahl der bis 2004 erfolgten »exonerations« ein Fehlerurteilsanteil von 2,3 % an den bundesweit erfassten Verurteilungen von 1973 bis 1989.³⁶ Bei Beschränkung auf die Untergruppe der 1982 bis 1989 verhängten Todesstrafen, denen erstens ein »rape-murder«-Vorwurf zugrunde lag und bei denen zweitens auch auswertbares biologisches Spurenmaterial vorhanden war, lag der Anteil an Entlastungsentscheidungen – mit nach wissenschaftlichen Maßstäben objektiver, weil durch nachträgliche DNA-Analyse belegter Unschuld – bis 2006 bei 3,3 %.³⁷ Dabei ist allerdings zu berücksichtigen, dass Verurteilungen zur Todesstrafe durchaus häufig nachträglich in eine lebenslange Freiheitsstrafe umgewandelt werden. In diesen Fällen sinkt die

32 Vgl. <https://deathpenaltyinfo.org/innocence-and-death-penalty> (Aufruf v. 4.12.2018). Erfasst werden Fälle mit Entlastungsentscheidungen nach Todesstrafe (für diesbezügliche Untersuchungen grundlegend *Liebman/Fagan/West/Lloyd* Texas Law Rev. 78 [2000], 1839 [1846 ff.]).

33 Vgl. <https://www.innocenceproject.org/cases/> (Aufruf v. 4.12.2018). Erfasst werden Fälle mit Entlastungsentscheidungen, die – oft auf Betreiben des Projektes hin – durch nachträgliche DNA-Tests erreicht wurden.

34 Vgl. <https://www.law.umich.edu/special/exoneration/Pages/detailist.aspx> (Aufruf v. 4.12.2018). Zur Würdigung dieser Datenbank vgl. *Leo J. Contemporary Criminal Justice* 33 (2017), 82 (87 ff.); siehe auch die Voruntersuchung von *Gross/Jacoby/Matheson J. Criminal Law & Criminology* 95 (2005), 523 ff.; zu einem sich im Aufbau befindlichen Nachahmerprojekt in Australien vgl. *Dioso-Villa Flinders Law J.* 17 (2015), 163 (177 ff.).

35 Dazu beispielsweise *Gross/Jacoby/Matheson J. Criminal Law & Criminology* 95 (2005), 523 ff.; *Gross Ohio State J. Criminal Law* 14 (2017), 753 (764 ff.).

36 Vgl. *Gross/O'Brien J. Empirical Legal Studies* 5 (2008), 927 (944 ff.). Zu einem ähnlichen Befund (speziell für »murder«-Verurteilungen) gelangte *Poveda* im Rahmen einer lokalen Sonderauswertung der New Yorker Strafgefangenenstatistik (*Justice Quarterly* 18 [2001], 689 [696 f.]). Aus der Anzahl von Entlassungen wegen nachträglichen Freispruchs ergab sich, bezogen auf die betreffende Gesamtfallzahl, eine Fehlerurteilsquote von 1,4 %.

37 *Risinger J. Criminal Law & Criminology* 97 (2007), 761 (768 ff.).

vergleichsweise hohe Wahrscheinlichkeit, mit der unbegründete Todesstrafenurteile aufgedeckt und aufgeklärt werden, wegen des verringerten Entlastungsdrucks deutlich. Wird das größere Dunkelfeld dieser Konstellationen in statistische Modellberechnungen integriert, soll die Fehlerurteilsquote bei Todesstrafenurteilen insgesamt 4,1 % betragen.³⁸

Jenseits des Sonderbereichs der Todesstrafenurteile lässt sich die Verbreitung von Fehlerurteilen anhand von Hellfeldmaterial an sich gar nicht bestimmen. Möglich wäre dagegen eine Dunkelfeldmethodik, bei der in *unausgelesenen*, d. h. zufällig ausgewählten Verfahrensstichproben jene Urteile, die Fehlerurteilsmerkmale aufweisen, bestimmt und hochgerechnet werden (zur auch hier bestehenden methodischen Problematik oben I.4.). Relativ tragfähige Befunde wurden auf diese Weise aber bislang allein durch eine Arbeit erreicht, die bei einem archivierten Bestand von 715 Altverfahren (1973 bis 1987 mit rechtskräftig gewordenen Strafurteilen zu Tötungs- und Sexualdelikten) eine neuerliche forensische Überprüfung des oft noch vorhandenen Spurenmaterials durchführen konnte. Unter den Fällen mit nachträglicher DNA-Analyse war diese in ca. 15 % »supportive of exoneration«.³⁹ Die nochmalige Klassifizierung und Überprüfung der Fälle (unter Hinzuziehung ergänzender Verfahrens- und Akteninformation) ergab schließlich eine Fehlerurteilsrate von 11,6 %.⁴⁰ Da die untersuchten Verfahren als weitgehend repräsentativ für das damalige sexualstrafrechtliche Vorgehen eingeschätzt werden, sind die Befunde – für den fraglichen Zeitraum und Deliktsbereich – tendenziell generalisierbar.⁴¹ Sie gelten so als »the strongest empirical evidence to date that establishes the existence of a substantial proportion of wrongful convictions in a sample of general cases«.⁴² Insbesondere für das große Feld der *leichten und mittelschweren* Delinquenz haben derartige Befunde allerdings keine Aussagekraft.

Bei der Massenkriminalität ist nicht nur die Wahrscheinlichkeit, dass fehlerhafte institutionelle Verarbeitungsformen aufgedeckt und korrigiert werden, eine andere, sondern auch die Fehleranfälligkeit selbst, die ihrerseits geprägt wird durch u. a. institutionelle Prozesse, die Rechtslage und die typischen Beweissituationen. Allerdings ist umstritten und ungeklärt, ob die Zahlen hier höher⁴³ oder niedriger⁴⁴ liegen. Dies betrifft etwa die enormen Fallanteile, die in den USA im plea bargaining-Verfahren erledigt werden.⁴⁵

38 Vgl. *Gross/O'Brien/Huc/Kennedy* Proceedings of the National Academy of Sciences 111 (2014), 7230 ff. für den Zeitraum 1973 bis 2004.

39 *Roman/Walsh/Lachman/Yahner* Post-conviction DNA testing and wrongful conviction, 2012, S. 28 ff.; vgl. hierzu ferner *Gross/O'Brien* J. Empirical Legal Studies 5 (2008), 927 (936 f.); *Gross* in: Huff/Killias (Hrsg.), *Wrongful conviction and miscarriages of justice*, 2013, S. 45 (55 f.).

40 *Walsh/Hussemann/Flynn/Yahner/Golian*, Estimating the prevalence of wrongful convictions, 2017.

41 Dass die abgeurteilten Verfahren in jener Zeit bundesweit recht homogen zusammengesetzt waren (gemessen an justizstatistisch verfügbaren Merkmalen), spricht nach *Walsh/Hussemann/Flynn/Yahner/Golian* (Fn 40) für eine vorsichtige Übertragbarkeit über den fraglichen Bundesstaat (Virginia) hinaus.

42 *Zalman/Grunewald* Texas A&M Law Rev. 3 (2015), 189 (237 Fn 324).

43 Wegen der geringeren Sorgfalt bei geringerer Vorwurfsschwere; entsprechende Hinweise für die Schweiz berichten *Gilliéron/Killias* FS Riklin, 2007, S. 379 (388 f.).

44 Bei schwerer Delinquenz steige der Erfolgsdruck für den Rechtsstab und damit auch die Fehlergefahr. Vgl. zur Debatte etwa *Gould/Leo* J. Criminal Law & Criminology 100 (2010), 825 (836).

45 Bei den Entscheidungen der U.S. District Courts liegt der Anteil insgesamt – mit regionalen und deliktsbezogenen Unterschieden – bei nahezu 90 % (dazu und zur sonstigen diesbezüglichen Forschung etwa *Johnson/King/Spohn* Annual Rev. Law Social Science 12 [2016], 479 ff.; vgl. auch die Angaben bei *LaFave/Israel/King/Kerr* Criminal Procedure, 6. Aufl. 2017, S. 19, 1194).

Dass hierzu so gut wie gar keine Fehlerurteilsstudien vorliegen,⁴⁶ macht Aussagen zu einer deliktübergreifenden Gesamtfehlerurteilsquote unmöglich. Zwar werden zur Klärung dieser Problematik bisweilen indirekte Methoden eingesetzt, doch sind diese zur Prävalenz-erfassung ungeeignet. Dies betrifft aus den oben (I.4.) genannten Gründen vor allem die vorliegenden Justizpraktiker-Surveys⁴⁷ und Interviews mit verurteilten Personen (insbesondere Strafgefangenen).⁴⁸ Bisweilen wird dennoch versucht, die unterschiedlichen Anhaltspunkte integrierend zu würdigen und in eine Gesamtschätzung zu überführen. Danach liege die US-amerikanische Fehlerurteilsprävalenz delikt- und verfahrenübergreifend bei 0,5 bis 1,0 % bzw. bei 1 bis 5 %.⁴⁹ Wirklich verlässlich sind diese Werte aber nicht. Sie erfahren deshalb mitunter auch Widerspruch von Autoren, die – freilich ebenfalls auf der Basis auch ungesicherter Annahmen – zu deutlich niedrigeren Raten gelangen.⁵⁰

2. Identifizierung der wichtigsten Fehlerurteilsquellen

In der neueren US-Forschung wird die Prävalenzfrage zunehmend durch die Untersuchung von Fehlerurteilsursachen verdrängt. Allerdings weisen die verfügbaren Analyseverfahren auch in diesem Bereich allesamt spezifische Schwachstellen auf.⁵¹ Den forschungspraktisch dominierenden Ansatz stellen induktiv vorgehende »aggregated case studies« dar. Dabei werden die in den oben (II.1.) genannten Datenbanken registrierten Strafverfahren ausgewertet und die Gründe für die jeweilige Fehlerurteilsentstehung identifiziert. Die Zusammenfassung zu Fehlerkategorien macht sodann deren statistische Verteilung (bezogen auf den gesamten ausgewerteten Fallbestand) deutlich.⁵² Jene Fehlerkategorien, die

46 Für eine Ausnahme vgl. die Auswertung von insgesamt 85 Fällen durch *Gross* Ohio State J. Criminal Law 14 (2017), 753 (776 ff.); *ders.* Boston Univ. Law Rev. 98 (2018), 999 (1000 ff.).

47 Bei *Huff/Rattner/Sagarin* Crime and Delinquency 32 (1986), 518 (522 f.) sehen 72 % der Befragten die Fehlerurteilsquote bei < 1 %. In den Befragungen von *Ramsay/Frank* Crime and Delinquency 53 (2007), 436 ff. sowie *Zalman/Smith/Kiger* Justice Quarterly 25 (2008), 72 ff. gehen die meisten Befragten von 0,5 % im eigenen Gerichtsbezirk und von 0,5 bis 5,0 % in den USA insgesamt aus. Vgl. auch die Befragungen zur Häufigkeit von Divergenzen zwischen den Entscheidungen und den Falleinschätzungen der dem Verfahren jeweils vorsitzenden Richter durch *Spencer* (J. Empirical Legal Studies 4 [2007], 305 ff.): Juryentscheidungen danach mit einer Wahrscheinlichkeit von 83 bis 89 % korrekt. Allerdings bleibt bei den hier zugrundeliegenden Befragungen unklar, ob die Richter wirklich wegen »tatsächlicher Unschuld« und nicht nur aus rechtlichen Gründen (»Prozessordnungswidrigkeit des Schuldbeweises«) ein anderes Ergebnis als das Urteil der Jury präferierten; i. Ü. bleiben Fälle, in denen Jury und Richter gemeinsam irren, von vornherein unberücksichtigt.

48 Diese Befragungen enthalten diverse reliabilitätssichernde Elemente und werden um statistische Verfahren zur Falschangabenschätzung ergänzt. So *Loeffler/Hyatt/Ridgeway* J. Quantitative Criminology 2018 (<https://link.springer.com/article/10.1007%2Fs10940-018-9381-1>): Fehlerurteilsquote von 6 bis 8 %. Bei *Poveda* (Justice Quarterly 18 [2001], 689 [698 ff.]) führte dieses Verfahren sogar zu einer Rate von 15 %.

49 Vgl. dazu *Zalman* Criminal Law Bulletin 48 (2012), 221 ff. bzw. *Gross* in: *Huff/Killias* (Fn 39), S. 45 (57). Für die Niederlande schätzt *Derksen* (Onschuldig Vast, 2016) auf der Grundlage von Interviews und diversen anderen Hinweisen eine Quote von 5 bis 10 %, wobei die Werte bei Vielfachtätern und bei schwerer Delinquenz noch deutlich höher liegen sollen. Diese Angaben sind als hysterisierend und spekulativ zurückgewiesen worden (vgl. *Merckelbach/Otgaar* Nederlands Juristenblad 2017, 450 ff.; zum Fehlen niederländischer Daten auch *Brants* University of Cincinnati Law Rev. 80 [2012], 1069 [1071 f.]).

50 Vgl. zuletzt *Cassell* Ariz. L. Rev. 60 (2018), 815 (826 ff.): 0,016 bis 0,062 %; für North Carolina *Thomas III* Ariz. L. Rev. 60 (2018), 865 (872 ff.): 0,125 bis 0,5 %.

51 Näher dazu die Übersichtsdarstellung bei *Cole/Thompson* in: *Huff/Killias* (Fn 39), S. 111 (115 ff.). Besondere Probleme werden erneut jene Arbeiten auf, in denen Justizpraktiker nach ihrer Auffassung zu Fehlerurteilsursachen befragt werden (für ein Beispiel vgl. *Jiahong/Ran* University of Cincinnati Law Rev. 80 [2012], 1277 [1282 ff.] für China).

52 Im Grunde wurde bereits in der klassischen Studie von *Peters* (Fn 1) in dieser Art verfahren.

gehäuft bzw. in besonders vielen Fällen auftreten (Fehlerschwerpunkte),⁵³ gelten hierbei als »immediate correlates of wrongful convictions«⁵⁴ und als *Risikofaktoren* – d. h. als Fehlerformen, denen das Potenzial, Fehlerurteile hervorzurufen, in stärkerem Maß innewohnt als anderen Fehlerarten. Die meisten bislang vorliegenden Studien haben dabei weitgehend übereinstimmende prozessuale Problemlagen festgestellt, so dass sich – auch wenn die Rangfolge der Fehlerurteilsrelevanz der Fehlerkategorien zwischen den verschiedenen Deliktsbereichen variiert⁵⁵ – ein regelrechter Fehlerkanon herauskristallisiert hat:⁵⁶ Unzulänglichkeiten im Bereich Kriminaltechnik/Forensic Sciences; Falschidentifizierung durch Augenzeugen; Falschaussagen/-beschuldigungen durch Zeugen und Tatbeteiligte; Falschgeständnisse; Verteidigungsfehler; Fehlverhalten von Polizei- und Justizpersonal; Einsatz unzuverlässiger Informanten.

Allerdings ist bei der Deutung dieser Befunde Zurückhaltung geboten. Zunächst macht sich auch in der Fehleranalyse der Umstand bemerkbar, dass das Untersuchungsmaterial aus dem Fehlerurteilshellfeld stammt und daher selektiert ist (oben I.3.).⁵⁷ In den ausgewerteten Fallsammlungen sind nämlich jene Verfahren stärker vertreten, bei denen die Bereitschaft zur Fehlersuche gesteigert ist und solche Fehler auftreten, die gut beweisbar und für gängige Fehlermachungsverfahren eher zugänglich sind – etwa Delikte mit biologischen Spuren und der Möglichkeit eines nachträglichen DNA-Tests. Andere Verfahrensbereiche – und damit auch die dort auftretenden Fehlerquellen – sind in den Fallsammlungen dagegen notorisch unterrepräsentiert.⁵⁸ Dies betrifft insbesondere auch Entscheidungen auf Basis eines *plea bargaining*, bei denen es vergleichsweise selten zu einer Wiederaufnahme kommt – weswegen beispielsweise der hierfür charakteristische Einwilligungs- und Geständnisdruck⁵⁹ in den genannten Fehlerschwerpunkten fehlt.⁶⁰ Insofern erlauben die vorliegenden Studien nur den Schluss, dass die genannten Fehlerformen überhaupt eine Rolle spielen. Über die Bedeutung dieser Rolle in der Gesamtheit der Fehlerurteile lässt sich hingegen keine Aussage treffen.⁶¹

53 Die ausführlichsten und aktuellsten Auswertungen dieser Art sind auf den in Fn 32 bis 34 genannten Internetseiten zu finden. Vgl. aus den übrigen Arbeiten, die sich auf diese Methode stützen, etwa die klassische Studie von *Bedau/Radelet* *Stanford Law Rev.* 40 (1987), 21 (56 ff.); ferner *Gross/Jacoby/Matthewson* *J. Criminal Law & Criminology* 95 (2005), 523 ff.; *Gould* *The innocence commission*, 2007, S. 74 f.; *New York State Bar Association's Task Force on Wrongful Convictions* Final Report, 2009; *Garrett* *Convicting the innocent*, 2011, S. 178 ff.; *ders.* in: Medwed (Hrsg.), *Wrongful convictions and the DNA revolution*, 2017, S. 40 (43 ff.).

54 *Bonventre/Norris/West* in: Redlich/Acker/Norris/Bonventre (Hrsg.), *Examining wrongful convictions*, 2014, S. 301 (302).

55 Dazu beispielsweise *Gross* *Ohio State J. Criminal Law* 14 (2017), 753 (771 ff.).

56 Vgl. dazu auch die Übersichten bei *Gould/Leo* *J. Criminal Law & Criminology* 100 (2010), 825 (841 ff.) und *Huff* in: Kapardis/Farrington (Hrsg.), *The psychology of crime, policing and courts*, 2016, S. 175 (177 ff.).

57 Vgl. *Gross* *Ohio State J. Criminal Law* 14 (2017), 753 (769 f., 773 f.).

58 Zum Problem unentdeckter Fehlerschwerpunkte eingehend auch *Zalman/Larson* *Albany Law Rev.* 79 (2016), 941 (952 ff., 982 ff.) unter Hinweis auf den Aspekt der unprofessionellen polizeilichen Gesamtermittlung.

59 Mit experimentellen Methoden, die eine Kontrollgruppentestung möglich machen (»cheating paradigm«), konnte gezeigt werden, dass das Anbieten eines Deals die Wahrscheinlichkeit sowohl von falschen als auch von wahren Geständnissen erhöht. Vgl. *Russano/Meissner/Narchet/Kassin* *Psychological Science* 16 (2005), 481 ff.

60 Vgl. aber erstes Fallmaterial bei *Gross* *Ohio State J. Criminal Law* 14 (2017), 753 (776 ff.).

61 Die besagten Arbeiten sind nicht frei von technischen Schwächen. Beispielsweise wird bei der Bildung von Fehlergruppen in der Regel nur berücksichtigt, dass es sich um einen »contributing factor« handelt, ohne dabei nach dem ganz unterschiedlichen Gewicht zu differenzieren, mit dem dieser Faktor zur Fehlerurteilung beigetragen hat. Ohnehin beruht die Bestimmung der fallrelevanten Fehlerurteilsursachen auf

Als Risikofaktor können die genannten Fehler ohnehin nur dann interpretiert werden, wenn sie in Fehlerurteilsverfahren wirklich stärker auftreten als in vergleichbaren Fällen, die richtig entschieden wurden.⁶² Feststellbar ist dies lediglich in Studien mit einem Vergleichsgruppendesign, das einen hohen Methodenaufwand verlangt⁶³ und deswegen bislang nur sehr selten eingesetzt wird. Eine dieser Studien verglich 105 Todesstrafenurteile (seit 1973), bei denen eine nachträgliche Entlastung vorlag, mit einer Zufallsauswahl vollstreckter Fälle (137 von 885) und wies auf diese Weise nach, dass die Fehlerurteilsgefahr bei schwieriger Beweislage steigt.⁶⁴ Für die hiesige Fragestellung unmittelbar relevant ist aber eigentlich nur eine Vergleichsstudie anhand von 260 Fehlerurteilsfällen (Unschuld nachträglich festgestellt) und 200 ähnlichen Verfahren, in denen es in oder schon vor dem Urteil zu einer Entlastung des Angeklagten gekommen war (jeweils Vorwürfe wegen eines Gewaltdelikts nach 1980). Hierbei bestätigte sich im Wesentlichen die Fehlerurteilsrelevanz der oben genannten Fehlergruppen – dies aber mit einigen bemerkenswerten Ausnahmen (Falschgeständnis, Informanteneinsatz und Polizeifehler).⁶⁵

III. Übertragbarkeit der Befunde

Eine unmittelbare Übertragung der US-amerikanischen Prävalenz- und Ursachenbefunde auf andere Rechtsordnungen ist nicht ohne Abstriche möglich. Dies wird durch die Ergebnisse nahegelegt, die in vereinzelt Studien aus diversen anderen Ländern erzielt worden sind und die die eben genannten Fehlerschwerpunkte (II.2.) nur eingeschränkt replizieren.⁶⁶ Von den neueren deutschen Arbeiten wird – wenngleich auf der Basis von sehr kleinen Fehlerurteilsansammlungen – ebenfalls die Bedeutung ganz anderer Fehlerkategorien dokumentiert, namentlich nicht erkannte psychische Erkrankung, Falschbeschuldigung sowie fehlender oder falsch erhobener Sachbeweis.⁶⁷ Letztlich sind solche Befunddifferenzen, selbst wenn sie teilweise auch mit den verschiedenen methodischen Herangehensweisen erklärt werden müssen, allerdings durchaus erwartungsgemäß. Die jeweiligen rechtlichen und tatsächlichen Prozessbedingungen weisen nämlich derart

einer subjektiven Bewertung der analysierenden Personen. Bei nicht wenigen Fällen, die in verschiedenen Studien von verschiedenen Kodierern ausgewertet wurden, hat dies eine uneinheitliche Fehlerkategorisierung zur Folge gehabt (vgl. <https://nij.gov/journals/279/Pages/wrongful-convictions-and-dna-exonerations.aspx#noteReference3> [Aufruf v. 4.12.2018]).

62 Zum Folgenden *Gross/O'Brien* J. Empirical Legal Studies 5 (2008), 927 (936 ff.).

63 Näher *Köbel* in: *Barton/Dubelaar/Köbel/Lindemann* (Fn 4), S. 45 f.

64 *Gross/O'Brien* J. Empirical Legal Studies 5 (2008), 927 (947 ff.). Ähnlich waren das Vorgehen und die Ergebnisse der etwas kleineren Studie von *Harmon* Justice Quarterly 18 (2001), 949 ff. (Bedeutung von Falschaussagen für Fehlerurteile). Siehe ergänzend auch *Garrett* Columbia Law Rev. 108 (2008), 55 ff.

65 *Gould/Carrano/Leo/Young* Predicting erroneous convictions: A social science approach to miscarriages of justice, 2012; *Gould/Carrano/Leo/Hail-Jares* in: *Zalman/Carrano* (Hrsg.), Wrongful conviction and criminal justice reform, 2014, S. 73 ff.

66 Vgl. für die Schweiz *Gilliéron/Killias* FS Riklin, S. 387 ff.; für Australien *Dioso-Villa* Flinders Law J. 17 (2015), 163 (181 ff.); für Norwegen *Stridbeck/Magnussen* Criminal Law Quarterly 58 (2012), 267 (279 ff.) und für China *Jiahong/Ran* University of Cincinnati Law Rev. 80 (2012), 1277 (1286 ff.); siehe ergänzend auch für Schweden *Hellqvist* Bergen J. Criminal Law and Criminal Justice 5 (2017), 133 (141 ff.). Bei solchen Vergleichen ist indes zu berücksichtigen, dass einige der genannten Fehlerkategorien sehr allgemein gehalten sind und ganz unterschiedliche Konstellationen einschließen (vgl. *Gross* Ohio State J. Criminal Law 14 [2017], 753 [775]).

67 *Dunkel* (Fn 5), S. 190 ff.; *Leuschner*, in: *Boers/Schaerff* (Hrsg.), Kriminologische Welt in Bewegung, 2018, S. 497 (501).

unterschiedliche Strukturen, Stärken und Schwächen auf, dass eine ähnliche Fehlurteilsanfälligkeit unwahrscheinlich und eine internationale Generalisierung von Forschungsergebnissen daher ausgeschlossen ist:

- Dies beginnt bereits damit, dass die summarischen, d. h. abgekürzten Verfahren im US- und im deutschen Verfahren (Absprachen, Strafbefehle) besonderen, teils gegensätzlichen Rahmensetzungen unterliegen. Dies beeinflusst die jeweiligen Falschentscheidungen unter den so abgeschlossenen Fällen,⁶⁸ hat aber auch erhebliche Auswirkungen auf die Selektion derjenigen Prozesse, in denen es noch zu einer Gerichtsverhandlung und einem dortigen Fehlurteil kommen kann. So ist etwa das staatsanwaltliche Ermessen beim exzessiv angewendeten⁶⁹ plea bargaining in den USA (bei pretrial-Verhandlungen) deutlich weiter als in den stärker formalisierten und regulierten Verständigungen der StPO. Auch ist der Richter in den USA nicht in die pretrial-Verständigungen involviert.⁷⁰
- Das Juryverfahren in den USA beruht auf der Annahme, dass Anklage und Verteidigung in der wechselseitigen Präsentation und Infragestellung von Fallversionen und Beweismaterial die Grundlage für eine wahrheitsnahe Fallrekonstruktion legen. Die Fehlervulnerabilität eines solchen Prozesstyps besteht darin, dass die vorausgesetzte Waffengleichheit angesichts oft limitierter Verteidigungskapazitäten eingeschränkt und der gebotene Beweisstoff durch die jeweiligen Anklage- oder Verteidigungsinteressen parteilich eingefärbt ist.⁷¹ Demgegenüber sind die institutionellen Akteure (Polizei, Staatsanwaltschaft, Gutachter, Gericht) im inquisitorischen Verfahren Deutschlands de jure zur Neutralität bei der Fallarbeit verpflichtet.⁷² Davon erwartet man zumindest im Hauptverfahren ein geringeres Fehlerniveau, wenngleich dieser »Objektivitätsvorsprung« aus mehreren Gründen relativiert werden muss (Aktenkenntnis und so eingeschränkte Unvoreingenommenheit der justiziellen Beteiligten;⁷³ tendenziell großzügigere Berücksichtigungsfähigkeit von rechtswidrig erlangtem Beweismaterial). Für die Ermittlungsphase ist ohnehin festzuhalten, dass sie

68 In der Schweiz bildet die abgekürzte Beweisführung, die bei Bagatelldelikten im Vorfeld von Strafbefehlserlassen kennzeichnend ist, einen Fehlerschwerpunkt (vgl. die Befunde von *Gillieron/Killias* FS Riklin, S. 388 ff.). Ein Potenzial falscher Verurteilungen wird bei summarischen Verfahren ferner darin gesehen, dass die bei streitiger Verhandlung drohende Strafe die Beschuldigten zum »Einlenken« drängt (so etwa *Jehle* Forensische Psychiatrie, Psychologie, Kriminologie 7 [2013], 220 [224 f.]).

69 Dazu Fn 45.

70 Vgl. jedoch zur hier unterschiedlichen Gesetzeslage in den Bundesstaaten etwa *LaFave/Israel/King/Kerr* (Fn 45), S. 1241.

71 Näher etwa *Brants* University of Cincinnati Law Rev. 80 (2012), 1069 (1078 ff.) unter Hinweis auf konkrete Aspekte (Gefahr parteilicher Gutachter, die sich an Interessen der Auftraggeber orientieren; Gefahr einer unkritischen Haltung der Staatsanwaltschaft gegenüber ihren eigenen Zeugen und Sachverständigen).

72 Dazu unter Hinweis auf diverse Konsequenzen (z. B. weniger aggressive Polizeivernehmungen; fehlende Verbindlichkeit eines Geständnisses) eingehend *Killias* in: Huff/Killias (Fn 39), S. 61 ff.

73 Neue Informationen, die dem Vorwissen widersprechen, werden systematisch ausgeblendet und unterschätzt, aktenkompatible Informationen hingegen systematisch überschätzt (zu den Studien näher *Schünemann* StV 2000, 159 [160 ff.]; siehe zu diesem Bias auch *Singelstein* StV 2016, 830 ff.). Ob speziell das Wissen um (einschlägige) Vorstrafen die Fallbeurteilung beeinflusst (und die Wahrscheinlichkeit der Schuldzuschreibung bei ansonsten gleicher Informationslage erhöht), ist unklar (zu widersprüchlichen Befunden bei US- und Schweizer Studien vgl. *Oswald* in: *Oswald/Bieneck/Hupfeld-Heinemann* [Hrsg.], *Social Psychology of Punishment of Crime*, 2009, S. 364 ff.; hierzu auch *Schlotthauer/Yundina* R&P 34 [2016], 43 [45 f.]).

hierzulande von Polizei und Staatsanwaltschaft dominiert und realiter oft aus deren parteilichem Selbstverständnis heraus betrieben wird.⁷⁴

- Aufgrund der herausgehobenen Stellung des Gerichts in Deutschland stellt das Unterlassen einer sachlich gebotenen Beweiserhebung weniger eine parteiseitige Fehlerquelle dar (wie im adversatorischen US-Verfahren) als eine solche des Gerichts. Überhaupt ist das deutsche, wie jedes inquisitorische System, abhängig von der Integrität und der Performance der Amtsermittler, so dass dessen Fehlerurteilsanfälligkeit aus hier ggf. bestehenden Mängeln erwächst. Dies gilt besonders auch deshalb, weil die Verteidigung dies nur bedingt kompensieren kann, da sie sich angesichts ihrer Rolle und ihrer regelmäßig geringeren Ausstattung nicht in der US-typischen Weise in der Fallkonstruktion engagiert.⁷⁵ Überhaupt besteht die Bedeutung der Verteidigung in Deutschland weniger in der Durchführung einer Gegenermittlung, sondern vor allem darin, auf die Regeleinhaltung durch Staatsanwaltschaft und Gericht zu achten. Verteidigungsversagen weist daher vermutlich eine geringere Fehlerurteilsrelevanz als in den USA auf.⁷⁶

IV. Konsequenzen und Perspektiven für Deutschland

1. Zur Forschungsnotwendigkeit

Auch wenn die Übertragbarkeit der oben referierten US-Befunde auf deutsche Verhältnisse ganz grundsätzlich bezweifelt werden muss und das deutsche Strafverfahren infolge seiner Orientierung an der materiellen Wahrheit für die Fehlerurteilsvermeidung vermutlich etwas besser gerüstet ist, bleibt die Annahme einer geringeren Fehlerurteilsanfälligkeit spekulativ.⁷⁷ Es ist jedenfalls nicht ausgeschlossen, dass die systemimmanenten Schwächen des deutschen Verfahrens zu einer durchaus relevanten Anzahl von Falschverurteilungen führen.⁷⁸ Dies liegt – auch angesichts erster Untersuchungen in anderen Staaten⁷⁹ – sogar nahe. Aufschlussreich dafür sind Erfahrungen aus einigen europäischen Ländern, in denen spezialisierte Institutionen eingerichtet worden sind, die an der Herbeiführung von ggf. erforderlichen Korrekturentscheidungen systematisch mitwirken, wie etwa in England und Wales sowie in Schottland und Norwegen.⁸⁰ Diese Kommissionen

⁷⁴ Dazu *Zalman/Grunewaldt* Texas A&M Law Rev. 3 (2015), 189 (220 f.); *Eisenberg/Köbel* Kriminologie, § 27 Rn 21 ff., § 28 Rn 8 ff.

⁷⁵ Vgl. dazu etwa *Brants* University of Cincinnati Law Rev. 80 (2012), 1069 (1079 ff.). Objektivitätsdefizite bei den Amtsermittlern (etwa ein polizeilicher Tunnelblick oder eine gewisse richterliche Voreingenommenheit nach der Vorbefassung mit den Ermittlungsakten) schlagen deshalb leichter in Fehlentscheidungen um (vgl. auch *Zalman/Grunewaldt* Texas A&M Law Rev. 3 [2015], 189 [224]).

⁷⁶ Vgl. *Huff* in: Kapardis/Farrington (Fn 56), S. 181; vgl. ferner *Killias* in: Huff/Killias (Fn 39), S. 61 ff.

⁷⁷ Vgl. auch *Killias* in: Lupária (Hrsg.), Understanding wrongful conviction, 2015, S. 57, 58 f.: Dass das »gefühlte« Fehlerurteilsaufkommen in Europa geringer als in den USA zu sein scheint, muss nicht an den Prozessstrukturen liegen, sondern könnte u. U. auch ein Fehleindruck sein und schlicht damit erklärt werden, dass das Spuren- und sonstige Beweismaterial nicht in der gleichen Weise wie in den USA konserviert und archiviert wird, was die Aufdeckung und den Nachweis von Fehlerurteilen in Europa erschwert.

⁷⁸ Vgl. *Grunewald* Albany Law Rev. 77 (2014), 1139 ff.

⁷⁹ Vgl. Fn 66.

⁸⁰ Für eine vergleichende Darstellung vgl. *Roach* in: Bruinsma/Weisburd (Hrsg.), Encyclopedia of criminology and criminal justice, Bd. 5, 2014, S. 2514 ff.; zur gescheiterten Institutionalisierung einer ähnlichen Kommission in den Niederlanden vgl. *Brants* University of Cincinnati Law Rev. 80 (2012), 1069 (1106 ff.).

untersuchen die ihnen vorgelegten Verfahren auf Fehlurteilsanzeichen und lösen ggf. eine abermalige Befassung durch die Justiz aus.⁸¹ Das von ihnen so bewirkte Aufkommen nachträglicher Freisprüche ist jeweils überschaubar, aber dennoch so *substanziell*, dass es die gesellschaftliche Virulenz der Fehlurteilsproblematik unterstreicht.⁸² Es handelt sich hierbei also wahrscheinlich auch in Deutschland um ein nicht zu vernachlässigendes Phänomen – insbesondere wenn man unabhängig von dessen Ausmaß die individuellen und gesellschaftlichen Folgen berücksichtigt.

Eine Fehlverurteilung bedeutet zunächst, dass der wahre Täter der Strafverfolgung entgangen ist.⁸³ Das spätere Bekanntwerden des Fehlers bringt somit die Tatgeschädigten und deren Umfeld in eine ausgesprochen ambivalente Situation.⁸⁴ Besonders gravierend sind indes die *individuellen Folgen* bei den fehlerhaft verurteilten Personen.⁸⁵ Bei diesen können zum einen Rechtsverletzungen – wie einige Fallstudien zeigen – insbesondere bei schweren Strafen eine ganz erhebliche Qualität annehmen.⁸⁶ Zum anderen reichen die tatsächlichen Folgen und Einbußen weit über diesen Bereich hinaus. Konkret geht es um Kosten und Einkommensverluste,⁸⁷ um verlorene Lebenszeit und Entfaltungsmöglichkeiten, um Beeinträchtigungen sozialer Beziehungen und Zwang zu unerwünschten sozialen Kontakten⁸⁸ und um vieles mehr.

Während des Verfahrens und im Strafvollzug entwickeln fehlerverurteilte Personen neben den allgemeinen Adaptionsstrategien auch Anpassungsformen, die spezifisch für die Situation einer unbegründeten Inhaftierung sind – unter anderem Rückzug, die Zurückweisung des Straftäterlabels und eine vertiefte Beschäftigung mit dem Freispruchsszenario.⁸⁹ Fehlerhaft Inhaftierte erfahren vielfach eine Umschreibung der Identität und hegen im Verlauf der Haft in der Regel zu min-

81 Sie befinden also nicht eigenständig über die Fälle, sondern veranlassen lediglich deren Wiederaufnahme (wobei sie allerdings die hierfür erforderlichen Ermittlungen führen). Demzufolge ist ihr Entscheidungskriterium auch nicht unbedingt die Unschuld der verurteilten Person, sondern die Wahrscheinlichkeit, mit der das fragliche Urteil vor Gericht (d. h. nach den justizeigenen Kriterien) korrigiert werden wird (kritisch daher *Naughton* *The innocent and the criminal justice system*, 2013, S. 162 ff.; näher dazu *Grunevald* *Albany Law Rev.* 77 [2014], 1139 [1183 ff.]; *King* in: *Redlich/Acker/Norris/Bonventre* [Fn 54], S. 217, 228).

82 So kam es beispielsweise in Norwegen zwischen 2004 und 2010 zu 120 Wiederaufnahmen (bei 1.184 der Kommission vorgelegten Fällen) – darunter zu 96 Freisprüchen sowie 17 Urteilen mit abgesenkter Strafe (*Stribeck/Magnussen* *Criminal Law Quarterly* 58 [2012], 267 [279 ff.]). Zu den Ergebnissen der CCRC-Fälle vgl. <https://ccrc.gov.uk/case-statistics/> (Aufruf v. 4.12.2018).

83 »Wrongful Liberty«. Vgl. dazu *Baumgartner/Grigg/Ramirez/Rose/Lucy* *Albany Law Review* 81/4 (2018), 1263.

84 Vgl. die Interviews bei *Jenkins* *Inter. Rev. of Victimol.* 20 (2014), 243 (249 ff.): Feindseligkeit und Misstrauen sowie Aufrechterhaltung des Verdachts gegenüber der (falsch) verurteilten Person nach deren Entlastung.

85 Hinzu kommen ggf. eintretende Rechtsverletzungen der mittelbar Betroffenen (Familie usw.) durch finanzielle Aufwendungen und psychische Beeinträchtigungen (vgl. *Hoffmann/Leuschner* [Fn 5], S. 72 f.).

86 Hierzu auch *Westervelt/Cook* in: *Huff/Killias* (Fn 39), S. 261 ff.; *Hoyle* in: *Bosworth/Hoyle/Zedner* (Hrsg.), *Changing Contours of Criminal Justice*, 2016, S. 270 ff.

87 Dazu anhand deutscher Fälle *Hoffmann/Leuschner* (Fn 5), S. 71 f., 81 ff.

88 Einschließlich Angriffen, Demütigungen und Einschüchterungen (vgl. *Grounds* *Canadian J. Crim and Crim Justice* 46 [2004], 165 [169 f.]).

89 *Campbell/Denov* *Canadian J. Crim and Crim Justice* 46 (2004), 139 (145). Die Nichtanerkennung der vorgeworfenen Straftat bringt i. Ü. vielfach Vollzugsnachteile und die Behandlung als »Leugner« mit sich (dazu *Campbell/Denov* *Canadian J. Crim and Crim Justice* 46 [2004], 139 [153]; *Grounds* *Canadian J. Crim and Crim Justice* 46 [2004], 165 [170]).

destens einem Zeitpunkt Suizidgedanken.⁹⁰ Überhaupt sind die (allerdings erst ansatzweise erforschten) psychischen Folgen oftmals erheblich, gerade bei unbegründeter Freiheitsstrafe. Durch die wechselseitige Verstärkung der mit einer Langzeithaftierung einerseits und einer Fehlverurteilung andererseits verbundenen Belastungen gehen die Folgen über die Persönlichkeitsveränderungen (etwa F62.0 ICD-10 und PTSD), die mit langer Freiheitentziehung ohnehin oft verbunden sind, vielfach deutlich hinaus.⁹¹ Dokumentiert wurden Schlafstörungen und Stimmungsschwankungen, Depressionen, Panikstörungen, paranoide Störungen sowie Alkohol- bzw. Drogenabhängigkeit.⁹² Die psychische Lage nach der Entlassung weist Ähnlichkeiten mit der Situation der Überlebenden von Katastrophen usw. auf, wobei die Betroffenen (im Unterschied zu anderen traumatisierten Menschen) seltener nach sozialer, psychologischer oder emotionaler Unterstützung suchen.⁹³ Dass die Entlassung häufig unvorbereitet geschieht,⁹⁴ trägt im Übrigen zu einer nicht selten überfordernden Konfrontation mit charakteristischen nachvollzuglichen Problemen bei, wie etwa Mittel-, Arbeits- und Wohnungslosigkeit, weggebrochene soziale Integration usw.⁹⁵

Eine eigenständige Komponente von Fehlurteilsfolgen besteht in der Beeinträchtigung des *Rechts-* und *Institutionenvertrauens*. Bei der Allgemeinheit wird dieses zwar erst bei einer Häufung von Falschverurteilungen nachhaltig leiden.⁹⁶ Bei den Geschädigten der fraglichen Tat hingegen tritt eine solche Wirkung regelmäßig schon durch die einzelne Falschentscheidung ein.⁹⁷ Besonders deutlich ist der Vertrauensverlust natürlich bei den fehlerhaft Verurteilten.⁹⁸ Verstärkt wird dies durch den Eindruck der Betroffenen, selbst noch nach der Korrektur der ungerechtfertigten Bestrafung durch die staatlichen Institutionen vernachlässigt zu werden. Der typischerweise bestehende Wunsch der Betroffenen, dass die ungerechte Behandlung staatlicherseits anerkannt und wiedergutmacht wird,⁹⁹ bleibt in den Augen der Betroffenen (oder auch objektiv) oft unerfüllt. Es kommt in der Regel weder zu einer Entschuldigung noch zu einer Verfolgung der im Ursprungsverfahren geschehenen (potenziell strafbaren) Aufklärungsfehler.¹⁰⁰ Wo eine finanzielle Kompensation erfolgt,¹⁰¹ wird diese von den Fehlverurteilten

90 *Campbell/Denov* Canadian J. Crim and Crim Justice 46 (2004), 139 (149, 157).

91 *Grounds* Canadian J. Crim and Crim Justice 46 (2004), 165 (168 f., 176, 178).

92 *Grounds* Canadian J. Crim and Crim Justice 46 (2004), 165 (169).

93 *Naughton* Brit. J. Criminol. 54 (2014), 1148 (1151).

94 *Grounds* Canadian J. Crim and Crim Justice 46 (2004), 165 (170 f.); *Hoffmann/Leuschner* (Fn 5), S. 88 f.

95 *Grounds* Canadian J. Crim and Crim Justice 46 (2004), 165 (171 ff.); *Naughton* Brit. J. Criminol. 54 (2014), 1148 (1151); *Hoffmann/Leuschner* (Fn 5), S. 63.

96 Erhebungen hierzu liegen indes nicht vor. Bei *de Keijser/de Lange/van Wilsem* *Punishment & Society* 16 (2014), 32 (44) zeigte sich allerdings eine Tendenz im »Verfahrenspublikum«, bei erheblichen Straftaten das Risiko von Fehlurteilen eher hinzunehmen und ggf. auch ein Absenken der Anforderungen an die Beweisdichte zu befürworten. Die (Sekundär-)Analyse von *Xiong/Greenleaf/Goldschmidt* Intern. J. Law, Crime and Justice 48 (2017), 14 (18 ff.) spricht für eine langfristig sinkende Akzeptanz der Blackstone Formel («it is better that ten guilty individuals escape than that one innocent suffer»).

97 Dazu *Naughton* Brit. J. Criminol. 54 (2018), 1148 (1152).

98 Zu deren besonderer Sensibilität gegenüber Ungerechtigkeiten und administrativen Fehlentscheidungen siehe *Campbell/Denov* Canadian J. Crim and Crim Justice 46 (2004), 139 (154).

99 *Campbell/Denov* Canadian J. Crim and Crim Justice 46 (2004), 139 (154).

100 *Naughton* Brit. J. Criminol. 54 (2014), 1148 (1149); *Hoffmann/Leuschner* (Fn 5), S. 86.

101 In UK und Wales in 99 % der Fälle (vgl. *Naughton* Brit. J. Criminol. 54 [2014], 1148 [1160]).

meist als unzulänglich und Affront angesehen. In Deutschland beruht dies nicht nur auf der Pauschalierung und bislang allzu geringen Zahlungshöhe, sondern auch auf deren Antragsabhängigkeit, durch die sich die fehlverurteilte Person als Bittsteller erlebt.¹⁰²

2. Gegenstände und Ausrichtung einer deutschen Fehler- und Fehlerurteilsforschung

Eine dezidierte Ausrichtung der Fehlerurteilsforschung auf die Prävalenzbestimmung ist angesichts der oben diskutierten Erhebungsprobleme (II.1.) keine Forschungsstrategie, die fruchtbare Erkenntnisse erwarten lässt. Sinnvoller ist es, die Forschungsressourcen auf die Identifizierung prozessualer Fehlerursachen zu konzentrieren. Zu bedenken ist dabei jedoch, dass eine Anbindung an den oben (bei I.) eingeführten Fehlerurteilsbegriff in dieser Hinsicht zu ganz erheblichen *Beschränkungen* führt. Unberücksichtigt bleiben hierdurch nämlich sämtliche Fehler, die während eines Strafverfahrens auftreten und die sich nicht nachteilig in der Tatsachengrundlage eines Strafurteils (oder auch gar nicht) auswirken. Dies betrifft etwa:

- Rechtsanwendungsfehler (Heranziehung einer nicht einschlägigen Strafnorm, Nichtberücksichtigung einer einschlägigen Strafnorm, Strafzumessungsfehler),
- Fehler, durch die es zu einer günstigeren Fallrekonstruktion und einer entsprechend vorteilhaften strafrechtlichen Bewertung für die betroffene Person kommt (etwa zu einem unbegründeten Freispruch),
- Fehler, die sich gar nicht in der Fallkonstruktion oder in einer falschen Sachverhaltsbildung niederschlagen (weil das fragliche Beweismaterial nicht berücksichtigt wird oder weil es allein um die allgemeine Verfahrensführung geht),
- Fehler, die sich nur in Zwischenentscheidungen auswirken, nicht aber im abschließenden Urteil (Verfahrenseinleitung, Anklage, Untersuchungshaftanordnung, Ermittlungseingriffe, erstinstanzliche Entscheidungen),
- und Fehler, durch die die Justizförmigkeit des Verfahrens beeinträchtigt wird, nicht aber die »materielle Wahrheit« der Fallrekonstruktion.¹⁰³

Der systematische Ausschluss dieser Fehler in der bisherigen Fehlerurteilsforschung ist dadurch gerechtfertigt, dass ihre Folgen tendenziell – aber freilich nicht in jedem Einzelfall – von geringerer Tragweite und von geringerer Eingriffsintensität im Hinblick auf die Grundrechte sind, als dies typischerweise bei den oben behandelten fehlerurteilserzeugenden Fehlern angenommen werden muss.¹⁰⁴ Soll die empirische Forschung indes dazu beitragen, in den Strafverfolgungsinstitutionen das Bewusstsein um die eigene Fehleranfälligkeit und die Bereitschaft zu einer entwickelten »Fehlerkultur« zu erhöhen, wäre eher eine Fehler- als eine Fehlerurteilsforschung erforderlich. Insofern ist eine diesbezügliche Bedarfs- und

102 Hoffmann/Leuschner (Fn 5), S. 74 f., 78.

103 Zur Unterscheidung von »factual innocence« und »legal innocence« (d. h. Verurteilungen nach schweren prozeduralen Fehlern, deren korrekte Berücksichtigung zu einem Freispruch hätte führen müssen) vgl. auch Risinger J. Criminal Law & Criminology 97 (2007), 761 (762 f.); Cassell Ariz. Law Rev. 60 (2018), 815 (819); für Deutschland dazu näher und differenzierend Barton FS Eisenberg, S. 15 (18 ff.).

104 Dazu Kölbel in: Barton/Dubelaar/Kölbel/Lindemann (Fn 4), S. 32 f.

Strategiediskussion angezeigt. Unabhängig von dieser Ausrichtungsfrage liegt es auf der Hand, dass die Ermittlung und Dokumentation von Fehler Schwerpunkten und/oder unmittelbaren Fehlerurteilsquellen nicht ausreichend sind. Man erfährt hierdurch nämlich nichts über deren *Entstehungszusammenhänge*. Dafür bedarf es Arbeiten, die sich dezidiert diesen »deeper causes«¹⁰⁵ widmen.

Weitgehend ausgeklammert bleiben kann dabei der Bereich des bewussten Fehlhandelns, der sich mit individuellen Interessen an einer Verfahrensobstruktion erklärt.¹⁰⁶ Solche Sachlagen sind auf individuelle Schädigungsintentionen zurückzuführen und daher kaum generalisierbar. Für eine Analyse der strukturellen Schwachstellen – und dies ist der Bereich, der rechtspolitischen Interventionen und der Entwicklung einer »Fehlerkultur« am ehesten zugänglich ist – sind sie kaum von Belang. Die empirische Forschung sollte sich deshalb primär auf die »handwerklichen Fehler« konzentrieren, die systemisch angelegt sind und daher ebenso systemisch angegangen werden können.¹⁰⁷

3. Entstehungszusammenhänge von Fehlern als Forschungsfeld

Die hiermit angesprochenen Hintergrundbedingungen werden in einigen Problematisierungslinien und disziplinären Zusammenhängen teilweise bereits nachdrücklich und in einer anschlussfähigen Weise diskutiert: Eine erste, eindeutig *psychologisch* fundierte Denkrichtung untersucht beispielsweise die limitierten »diagnostic capabilities« der Strafverfolgungsakteure bei der Bewertung von fragwürdigem Beweismaterial. Lügen, Irrtümer oder andere Fehlleistungen von Zeugen zu erkennen, ist ohne Expertenwissen von vornherein schwer.¹⁰⁸ Dies gilt gerade auch bei falschen Geständnissen.¹⁰⁹ Außerdem erhöhen Urteilsheuristiken, die von den Beteiligten gleichsam unreflektiert und ohne Überprüfung ihres Gültigkeitsanspruchs eingesetzt werden, wie auch kognitive Verzerrungen das Risiko, eine unzutreffende Vorstellung von den aufzuklärenden Ereignissen zu entwickeln. Solche zusammenhängenden und sich wechselseitig verstärkenden psychischen Prozesse – etwa die Neigung, einmal gebildete Annahmen zu stützen und durch eine hiervon geleitete Informationssuche und -interpretation zu bestätigen – können zu einem »Tunnelblick« führen, bei dem sich die fragliche Person (Richter, Sachverständiger, Ermittler) vorzeitig auf eine, ggf. auch

¹⁰⁵ Leo J. Contemporary Criminal Justice 21 (2005), 201 (213).

¹⁰⁶ Diese Variante (zu ihr etwa Killias in: Lupária [Fn 77], S. 62 ff.) liegt etwa in den von Turvey (Forensic fraud, 2013, S. 137 ff.) analysierten 100 Fällen von »forensic fraud« oder bei eigennützigem Falschsagen von »jailhouse informants« vor (speziell dazu näher Zimmerman in: Westervelt/Humphrey [Hrsg.], Wrongly convicted, 2001, S. 55 ff.; Turvey in: Turvey/Cooley [Hrsg.], Miscarriages of justice: Actual innocence, forensic evidence, and the law, 2014, S. 149 ff.; aus den Fallsammlungen vgl. etwa New York State Bar Associations Task Force [Fn 53], S. 6 f., 114 f.).

¹⁰⁷ Zur Differenzierung der intentionalen und nicht intentionalen Fehlleistungen auch Naughton (Fn 81), S. 27 ff.: »abortion of justice« und »miscarriage of justice«.

¹⁰⁸ Näher hierzu Simon In doubt, 2012, S. 150 ff., 168 ff.; speziell zu Fehlvorstellungen, denen norwegische und US-Richter bei der Bewertung von Augenzeugenaussagen unterliegen, vgl. die Daten bei Magnusen/Wise/Raja Psychology, Crime & Law 14 (2008), 177 ff.

¹⁰⁹ Zu den situations- und personenbezogenen Risikofaktoren, die deren Wahrscheinlichkeit erhöhen (z. B.: Minimierungs- und Maximierungstechniken der Polizei, lange Befragungsdauer, junges Alter, intellektuelle Beeinträchtigungen), vgl. überblicksartig Kassir/Drizin/Grisso/Gudjonsson/Leo/Redlich Law and Human Behavior 34 (2010), 3 ff.; Eisenberg/Kölbel Kriminologie, § 28 Rn 47 ff.; zur Problematik auch Fn 59.

falsche Fallversion festlegt¹¹⁰ und einseitig nur noch diese verfolgt (»forensic confirmation bias«).¹¹¹

Eine eher soziologisch ausgerichtete Perspektive bezieht sich auf die allgemeinere Forschung zu »organizational wrongdoing«¹¹² und betrachtet vor diesem Hintergrund die *organisationsspezifischen* Prozesse, die auf Fehlerurteile hinauslaufen können. Diese stellen gewissermaßen »organizational accidents« dar, zu denen es durch das Zusammentreffen oder die Verkettung von für sich genommen jeweils eher kleineren und folgenarmen Verstößen, Nachlässigkeiten oder Fehler-routinen bei verschiedenen Beteiligten kommt. Die oben erwähnten Fehler-schwerpunkte gelten danach nicht als eigenständige Ursachen, sondern als überschaubare, aber doch oft weichenstellende Schritte innerhalb typischer Entwicklungs-pfade, die schließlich in ein Fehlerurteil münden.¹¹³ Aus dieser Warte ist nach den organisationsinternen und -externen Bedingungen zu fragen, die in den Ermittlungspraktiken zu fehlerhaften Verdachtsfestlegungen drängen – etwa dem medialen Druck, den defizitären Ressourcen, den blind machenden Arbeitsbünd-nissen oder etwa auch nach jenen Ausprägungen der Polizei- und Justizkultur, die bei den amtlichen Strafrechtsakteuren eine »Verfolgerhaltung« und eine eher überführungs- als wahrheitsorientierte Handlungslogik bedingt oder zumindest fördert.¹¹⁴

Das Hauptinteresse einer dritten, vorwiegend juristischen Sichtweise gilt den Anteilen des *Rechts*. Bei den hier zugehörigen Arbeiten geht es darum, ob und wie bestimmte häufige Fehler durch das Normprogramm begründet, erleichtert oder begünstigt werden. Dazu bedarf es spezifischer, auf die einzelnen Fehler bezogener Untersuchungen, die sich um die Differenzierung der hintergründigen »Gefahrenquellen« bemühen.¹¹⁵ Ein Beispiel hierfür bietet die Täteridentifizierung durch Augenzeugen, die schon wegen der Wahrnehmungs- und Erinnerungs-probleme fehleranfällig ist,¹¹⁶ die aber ganz besonders unter suggestiven Einflüssen bei der Befragung und Gegenüberstellung leidet.¹¹⁷ In diesem Zusammen-

110 Für einen Überblick über die dem zugrundeliegenden kognitiven Verzerrungen siehe beispielsweise *Findley* in: Cutler (Hrsg.), *Conviction of the innocent*, 2012, S. 303 ff.; *Kassin/Dror/Kukucka* *Journal of Applied Research in Memory and Cognition* 2 (2013), 42 ff.; *Oswald/Wyler* in: Barton/Dubelaar/Kölbel/Lindemann (Fn 4), S. 103 ff.; speziell polizeibezogen *Singelstein* *StV* 2016, 830 ff.

111 Vgl. zu so entstehenden Fehlerurteilen etwa die qualitative Fallauswertung bei *Gould/Carrano/Leo/Young* (Fn 65), S. 72 ff., 84) für die USA und *Shiyuan* (*California Western Law Rev.* 54 [2018], 103 [113 ff.]) für China; ferner auch *Brants* in: Huff/Killias (Fn 39), S. 161, 170 ff. Speziell zur Frühphase des Tunnelblicks und den Fallbedingungen, die dessen Entstehung fördern, vgl. die Auswertung des bei Fn 65 beschriebenen Materials durch *Lowrey-Kinberg/Senn/Gould/Hail-Jares* *California Western Law Rev.* 54/1 (2018), Art. 2; *dies.* *Crime & Delinquency* 2018 (<https://doi.org/10.1177%2F0011128718793618>).

112 Grundlegend zu dieser Einbettung *Lofquist* in: Westervelt/Humphrey (Fn 106), S. 174 (175 ff.).

113 *Doyle* *Albany Law Rev.* 80 (2016), 1 ff.; vgl. auch *Gould/Leo* *J. Criminal Law & Criminology* 100 (2010), 825 (840 f.); *Dioso-Villa* *Flinders Law J.* 17 (2015), 163 (192 f.).

114 Zu entsprechenden Hinweisen aus der Polizeiforschung siehe *Eisenberg/Kölbel* *Kriminologie*, § 27 Rn 42 sowie speziell im hiesigen Zusammenhang *Crowder/Turvey* in: *Turvey/Cooley* (Fn 106), S. 49 ff.; zu Anhaltspunkten für Zusammenhänge zwischen der institutionellen Punitivität und der Fehlerurteilsentstehung vgl. *Gould/Carrano/Leo/Young* (Fn 65), S. 67.

115 Zusammenfassend zum Bestand der diesbezüglichen Forschung etwa *Garrett* (Fn 53), S. 8 ff., 246; siehe auch die Literaturübersicht bei *Leo* *J. Contemporary Criminal Justice* 33 (2017), 82 (90 f., 93); näher beispielsweise *Simon* (Fn 108) sowie die Beiträge in Cutler (Fn 110).

116 Die korrekten Identifizierungsraten sind daher auch unter günstigen Bedingungen meist gering (vgl. *Wells/Greathouse/Smalarz* in: Cutler [Fn 110], S. 167, 174 f. m. w. N.).

117 In den ausgewerteten Fehlerurteilsfällen war es oft zu beeinflussenden Hinweisen und Bemerkungen

hang wird von der US-Forschung immer wieder auch die Judikatur des Supreme Court problematisiert, der unter bestimmten (als fragwürdig geltenden) Voraussetzungen die Verwertung selbst von mangelhaft bzw. suggestiv gestalteter Wiedererkennungssprozeduren erlaubt.¹¹⁸ In Gestalt dieser justiziell eröffneten Spielräume für die Einführung problematischen Beweismaterials wird im Grunde also die Rechtslage als ein Einfallstor für fehlerurteilsfördernde Beweisführungen ausgemacht.

Ein vierter Diskursstrang befasst sich schließlich mit Fehlerurteilsfällen, bei deren Entstehung defizitäre kriminaltechnische Beweisbeiträge eine entscheidende Rolle gespielt haben.¹¹⁹ Dies betrifft zunächst Fehldeutungen durch die Gutachter,¹²⁰ insbesondere aber die fachlichen Schwächen mancher *Forensic Sciences*. Hier besteht beispielsweise oftmals das Problem, dass am Tatort aufgefundenene Spuren (Haare, Patronen, Biss-, Reifen- oder Fußabdrücke) in zahllosen Verfahren mit Material des Verdächtigen verglichen und dann ggf. als »Treffer« eingeordnet werden, ohne dass immer auch hinreichende Daten zur Häufigkeitsverteilung der als kennzeichnend geltenden Merkmale zur Verfügung stünden.¹²¹

Wünschenswert ist es, all diese Untersuchungen zu den »deeper causes« in einer doppelten Hinsicht *integrativ* auszugestalten. Erstens sollten sie mit der Fehlerschwerpunktforschung koordiniert sein und zielgerichtet die Hintergrundbedingungen gerade für jene Fehlerformen analysieren, die besonders gravierend und/oder besonders häufig sind.¹²² Zweitens muss berücksichtigt werden, dass die eben skizzierten Diskurslinien ihre »unit of analysis« in der Strafverfolgungspers-

durch Polizeibeamte und/oder zu suggestiv gestalteten Wiedererkennungsverfahren gekommen (vgl. *Garrett* [Fn 53], S. 54 ff.). Für einen allgemeinen Überblick über Einflussfaktoren von Personenidentifizierungen siehe beispielsweise *Sporer/Sauerland/Kocob* in: Bliesener/Lösel/Köhnken (Hrsg.), Lehrbuch Rechtspsychologie, 2014, S. 156 ff.

- 118 Die maßgebliche Leitentscheidung – die zeitlich noch vor der Herausbildung der hierzu inzwischen vorliegenden psychologischen Forschung erging – überträgt dem Richter die Einschätzung, ob das Verfahren suggestiv gestaltet war. Die Zeuenaussage darf aber auch bei Bejahung dieser Frage durchaus noch verwertet werden, wenn die Zeugenangaben vom Richter als zuverlässig eingeschätzt werden. Dafür gibt der Supreme Court einige (als vorwissenschaftlich geltende) Kriterien vor – u. a. wie sicher sich der Zeuge ist, wie seine ursprüngliche Beschreibung aussah, wieviel Zeit seit dem Ereignis vergangen ist usw. (für eine kritische Auseinandersetzung im Fehlerurteilskontext vgl. etwa *Garrett* [Fn 53], S. 53 f., 63 ff.; *Wells* [Fn 116], S. 169 ff.).
- 119 Die betreffenden Fehlerurteilsstudien zusammenfassend *Smit/Morgan/Lagnado* *Science and Justice* 58 (2018), 128 ff.; siehe auch <https://mij.gov/journals/279/Pages/wrongful-convictions-and-dna-exonerations.aspx#noteReference3> (Aufruf v. 4.12.2018); vgl. ferner die Aktenanalysen von *Garrett/Neufeld* *Virginia Law Rev.* 95 (2009), 1 ff. zu wissenschaftlich nicht gesicherten Gutachteraussagen in mündlichen Verhandlungen.
- 120 Zum Nachweis subjektiver Deutungsspielräume bei DNA-Analysen und zu deren Beeinflussung durch Vorannahmen der Sachverständigen vgl. *Dror/Hampikian* *Science and Justice* 51 (2011), 204 ff.
- 121 Eingehend *President's Council of Advisors on Science and Technology* *Forensic science in criminal courts: Ensuring scientific validity of feature-comparison methods*, 2016. Generell zu diversen Fehlerquellen in den Forensic Sciences vgl. *National Research Council* *Strengthening forensic science in the United States: A path forward*, 2009; Überblicksdarstellungen bei *Cole/Thompson* in: Huff/Killias (Fn 39), S. 120 ff.; *Hamirani/Dror/Morgan* *Albany Law Rev.* 81 (2017), 975 ff.; *Smit/Morgan/Lagnado* *Science and Justice* 58 (2018), 128 ff.
- 122 Hierzulande liegen zahlreiche Untersuchungen zu »strafprozessualen Fehlerquellen« vor – allerdings ohne systematische Bezugnahme auf die Fehlerurteilsproblematik (vgl. aus neuerer Zeit etwa *Neuhaas* *StV* 2015, 185 sowie die Aufsatzserie von *Deckers* *StraFo* 2015, 265; *Püschel* *StraFo* 2015, 269; *Barton* *StraFo* 2015, 315 und *Velten* *StraFo* 2015, 354; siehe ferner auch *Eisenberg* *R&P* 1/2019 und *ZJJ* 1/2019, jeweils im Erscheinen). Dass dabei aber auch die rechtstatsächlich relevanten Fehler diskutiert werden, ist ohne die im Text angemahnte Abstimmung mit der Fehlerschwerpunktforschung nicht gesichert.

son, der Justizorganisation, dem Rechtsinstitut oder der forensischen Disziplin verorten und deshalb in der Regel die jeweiligen psychologisch, organisationell, rechtlich oder fachwissenschaftlich angelegten Risikopotenziale aufzeigen. Was die fraglichen Arbeiten nicht erfassen, ist die Frage, wie die verschiedenen Fehlerquellen in einzelnen Verfahren interagieren und in welchen Prozessen sie sich in den verschiedenen Prozessphasen konkret entfalten.¹²³ Demonstriert wird vielmehr nur die »abstrakte Gefährlichkeit« von kognitiven Verzerrungen, Organisationskulturen und Regulierungs- oder Wissensdefiziten.¹²⁴ Deshalb bedarf es ergänzender *Fallstudien*, die die Realisierungsprozesse des jeweiligen problem-spezifischen Risikopotenzials durch qualitative Analysen exemplarischer Verfahren nachvollziehen.¹²⁵ Auf diese Weise kann sichtbar gemacht werden, wie bestimmte Fehler zu bestimmten Folgeentscheidungen bei der weiteren Informationssuche führen, etwa wie sich objektive Fehler in den Vorstellungen niederschlagen, die sich Entscheidungsträger vom Fallgeschehen machen.¹²⁶ Dies stellt eine wichtige Ergänzung der »Ursachenidentifizierung« und der »Ursachenerklärung« dar¹²⁷ – insbesondere wenn auch Verfahren untersucht werden, in denen die Fehler durch prozessuale Vorkehrungen aufgefangen bzw. neutralisiert werden.¹²⁸

V. Bilanz

Fehlurteile sind international als ein Forschungsfeld etabliert, das – ungeachtet der methodischen Probleme und Begrenzungen – zuletzt eine rapide Entwicklung erfahren hat. Dahinter bleibt der deutsche Forschungsstand immer deutlicher zurück. Dies ist ebenso wie das hiesige Desinteresse an jenem Thema weder hinnehmbar noch wird es dem Problemgehalt auch nur annähernd gerecht, wie nicht zuletzt die erheblichen Folgen für Geschädigte und Fehlverurteilte, aber auch für das Rechts- und Institutionenvertrauen vor Augen führen. Da sich diese für das deutsche Strafverfahren bestehende Lücke nicht durch Rückgriff auf ausländische Befunde schließen lässt, führt kein Weg an einem Ausbau der deutschen Fehler- und Fehlurteilsforschung vorbei. Dafür bedarf es des Aufbaus einer einschlägigen Falldatenbank und einer systematischen Analyse, die in ihrem Vorgehen nicht hinter die international vorhandenen Methodenerfahrungen zurückfällt. Der einzuhaltende Standard ist mit dem Einsatz komparativer Auswertungsverfahren (oben II.2.) und integrativer Forschungsstrategien (oben IV.3.) klar gesetzt.

123 Leo J. *Contemporary Criminal Justice* 21 (2005), 201 (210).

124 Zur Problematik vgl. *Poveda Justice Quarterly* 18 (2001), 689 (693 f.); *Cole/Thompson* in: *Huff/Killias* (Fn 39), S. 119 f.; *Bonventre/Norris/West* in: *Redlich/Acker/Norris/Bonventre* (Fn 54), S. 302.

125 Als Beispiele dafür siehe die Arbeiten von *Garrett* ([Fn 53], S. 14 ff.) für die USA; *Brants* (*University of Cincinnati Law Rev.* 80 [2012], 1069 [1090 ff.]) für die Niederlande und *Jiang* (*Wrongful convictions in China*, 2016, S. 42 ff.) für China.

126 Vgl. etwa *MacKillop/Vidmar Chicago-Kent Law Rev.* 90 (2015), 957 (969 ff.).

127 Näher *Zalman/Larson Albany Law Rev.* 79 (2016), 941 (971 ff.).

128 Eine Neutralisierung von »kontaminiertem Material« kann beispielsweise durch die Fallarbeit der anderen Prozessbeteiligten, durch Beweiswürdigungs- und Verwertbarkeitsvorgaben, aber auch durch Rechtsmittelverfahren erfolgen (speziell zu den keineswegs seltenen Interventionen von Rechtsmittelgerichten bei »misleading evidence« vgl. *Smit/Morgan/Lagnado Science and Justice* 58 [2018], 128 ff. am Beispiel von England und Wales). Beispielsweise führen keineswegs alle Falschidentifizierungen und Falschgeständnisse zu einer Verurteilung, vgl. *Gross Ohio State J. Criminal Law* 14 (2017), 753 (775) m. w. N.