

1 Psychiatrie, Gefährlichkeit und Prognose

Hans-Jörg Albrecht

1.1 Zur Einführung: Gefährlichkeit und gefährliche Prognosen

Dr. Tod wohnt in Texas. Sein bürgerlicher Name lautet James Grigson. Von Beruf ist Dr. Tod Psychiater. Texanische Geschworene, Staatsanwälte und Richter vertrauten ihm und er ist bekannt geworden wegen seiner Prognosen fortdauernder Gefährlichkeit (Marquart et al. 1989), die in Texas (aber auch in anderen amerikanischen Bundesstaaten) Voraussetzung für die Verhängung der Todesstrafe und für die Exekution von Menschen ist, die wegen Mordes ersten Grades von Geschworenen schuldig gesprochen wurden. Auf Texas sind bis zum 31.12.2010 464 der 1.234 für die USA insgesamt registrierten Exekutionen entfallen (Death Penalty Information Center 2010, 1). Für jede dieser Exekutionen war Voraussetzung die Aussage sachverständiger Personen, und es war häufig (nämlich in 140 Fällen) die Aussage von Dr. Tod, dass die Verurteilten, würden sie weiterleben, gefährlich wären. Dr. Tod ist wegen seiner eigenwilligen Prognosen, die in der Regel auf eine „100%ige Wahrscheinlichkeit“ schwerer Gewalt lauteten, aus der Amerikanischen Vereinigung für Psychiatrie im Übrigen ausgeschlossen worden (Marquart et al. 1989).

Der Tod als Folge gerichtlich angeordneter Exekutionen ist nicht nur aus Deutschland, sondern aus (fast ganz) Europa verschwunden. In Deutschland (und in Europa) hat sich dafür Dr. Sicherheit angesiedelt. Seine Prognosen sind Voraussetzung dafür, dass neben der Verurteilung zu Strafe, Sicherungsverwahrung oder die Unterbringung in der Psychiatrie angeordnet oder dass

Gefangene aus lebenslanger oder langer Freiheitsstrafe entlassen werden oder dort verbleiben. Unabhängig davon, ob wie in England/Wales eine unbestimmte Freiheitsstrafe (Imprisonment for Public Protection), die in strafende und sichernde Bestandteile aufgespalten ist, oder eine Sicherungsverwahrung wie in Deutschland oder in der Schweiz verhängt wurde, erfolgt eine Entlassung erst, wenn Gefährlichkeit auf der Grundlage eines Sachverständigenutachtens nicht mehr angenommen wird. Die Sprache der Psychiatrie und diejenige des Strafgesetzes unterscheiden sich und selbst innerhalb der jeweiligen Rahmen ist der Sprachgebrauch nicht unbedingt homogen (Nedopil 2006, 22), geht es um Gefahr, Gefährlichkeit und Wahrscheinlichkeit. Wenn von Gefahr, gar „höchster Gefahr“, Prognose und erhöhtem Prognosemaßstab (Mosbacher 2011, 230), Wahrscheinlichkeit oder Sicherheit die Rede ist, dann nehmen diese Begriffe aus der Perspektive von Juristen und Psychiatern unterschiedliche Bedeutungen an. Über solche Unterschiede in der Bedeutungszuschreibung und vor allem im jeweiligen Umgang mit Wahrscheinlichkeit und Risiko (Gigerenzer 2002) ist wenig bekannt.

Ein besonderes Problem resultiert bei der Feststellung besonderer Gewalttrisiken aus der Berufung auf eine besondere psychiatrische Expertise. Psychiatrisches Wissen soll es erlauben, vorherzusagen, wie hoch die Wahrscheinlichkeit (oder Gefahr) ist, dass Menschen (wieder) schwere Gewalthandlungen begehen. Hieraus ergibt sich eine Differenzierung und zwar zwischen Gewalt, die mit Irrsinn assoziiert ist, und einer solchen, die nicht aus Krankheit entsteht, sondern aus einer (freien) Entscheidung. Überzeugungen zu kausalen Zusammenhängen zwischen Gewalt und Irrsinn sind ebenso fest eingegraben wie Vorstellungen zu einem Gewohnheitsverbrecher, der gar nichts anderes mehr kann, als Verbrechen zu begehen.

Fälle, in denen Psychiater (oder Psychologen) das Auftreten von Gewalttaten nicht vorhergesagt haben, sind leicht zu finden. Dazu gibt es keine systematischen Untersuchungen, jedoch weist regelmäßige (und meistens ebenso beiläufige, weil Tötungen nämlich „in der Familie bleiben“¹) Berichterstattung in den Medien darauf hin, dass schwerer Gewalt zuweilen psychiatrische Beurteilungen vorhergehen, die die Gewalt eben nicht „vorhergesehen“ haben. Dass forensischen Expertisen seit Mitte der 1990er-Jahre in Deutschland immer mehr Aufmerksamkeit zuteil wird, ist offensichtlich nicht der Fehleranfälligkeit geschuldet, sondern einer zunehmenden Bedeutung bestimmter Sicherheiten (und Ängste) in sozialen, rechtlichen und politischen Diskursen, die mutmaßlich wiederum mit der Erosion gewohnter sozialer Ordnungen und Gewissheiten zusammenhängt (Sofsky 2006).

Die Debatten um die Gefährlichkeitsprognose lassen Entwicklungen erkennen, blickt man auf die forensischen Wissenschaften und die normativen, in

1 Spiegel Online 18.11.2010, „Familiendrama in München. Psychisch Kranker ersticht Vater vier Jahre nach Mutter.“ oder: Spiegel Online 13.04.2011, „Mord an Zwillingmädchen. Psychiater entließen labile Mutter nach Hause.“

Gesetzesreform und gerichtlichen Praktiken sichtbaren Ausformungen. Die Debatten verweisen auf eine regional sehr unterschiedliche Intensität mit Schwerpunkten in Nordamerika und Teilen Europas. Die damit zusammenhängenden Probleme werden in der forensischen, weniger in der kriminologischen Forschung und Literatur seit den 1990er-Jahren immer intensiver diskutiert. Die Rechtsprechung greift international die Probleme teilweise auf (mit auch hier regional signifikanten Unterschieden), ohne dass sich aber ihr Grundtenor, nämlich die Zulassung und Verwendung von Gefährlichkeitsprognosen in Entscheidungen über massive Eingriffe in das Freiheitsrecht, geändert hätte. Die Rechtsprechung legt an die Grundlagen der Gefährlichkeitsprognose offensichtlich andere Maßstäbe an als an andere Befunde, die sich auf eine wissenschaftliche (oder technische) Grundlage berufen.

1.2 Prognosen – ein altes Forschungsfeld der Kriminologie

Prognosen und darunter Gefährlichkeitsprognosen zählen einerseits zu den klassischen Fragestellungen der Kriminologie, andererseits repräsentieren sie heute ein wichtiges Kennzeichen des auf Sicherheit bedachten Strafrechts. Die Thematik der Gefährlichkeitsprognose ist verknüpft mit Theorien der Kriminalität, mit der Untersuchung des Rückfalls und ist dann verbunden vor allem auch mit Fragestellungen des Strafvollzugs und der Unterbringung in forensischen Einrichtungen, also mit den Möglichkeiten, durch gezielte Interventionen das Auftreten von (Rückfall-)Kriminalität zu beeinflussen.

Der Schwerpunkt der Prognoseforschung und der Entwicklung von Prognoseinstrumenten liegt in Nordamerika. Dort wurden bereits recht früh, ab den 1970er-Jahren, Probleme der Gefährlichkeitsprognose aus rechtlicher, ethischer und methodischer Sicht aufgegriffen. In diesen Forschungsansätzen ging es auch um die heute noch wirksame Frage, wie Gefährlichkeit überhaupt definiert und festgelegt werden kann, und um eine kritische Betrachtung insbesondere des Umgangs mit dem Konzept der Gefährlichkeit durch Gesetzgeber und Gerichte. Den Ergebnissen wurde dabei eine zirkuläre Argumentation attestiert (Shah 1978). Im Zentrum stand die Feststellung, dass Gefährlichkeit ein relatives Konzept sei und ferner lediglich aus einem komparativen Zugang verstanden werden könne (Auerhahn 2006). Sodann wurde darauf hingewiesen, dass Gefährlichkeit eine kriminalpolitische Kategorie darstelle und deshalb nicht in den Kompetenzbereich psychiatrischer klinischer Diagnosen fallen könne (Thornberry u. Jacoby 1979). Kritische Aufmerksamkeit wurde zunächst vor allem gutachterlichen Stellungnahmen zur Frage der Gefährlichkeit von Psychiatriepatienten zuteil. Ausgangspunkt war dabei die Kritik an der sogenannten klinischen Prognose, die im Vergleich zu aktuarischen bzw. statistischen Prognoseverfahren als wenig treffsicher angesehen wurde (Monahan 2003, 16ff.).

Die Debatte zu Gefährlichkeitsprognosen wurde in den USA insbesondere befeuert durch eine Feststellung Monahans, der 1981 Untersuchungen zur Pro-

gnose von Gewalt in dem Satz zusammenfasste, klinische (psychiatrische und psychologische) Prognosen trafen in höchstens einem von drei Fällen zu. Es gehe um Personen, für die eine Geisteskrankheit festgestellt worden sei und für die eine hohe Belastung mit Gewalt in der Vergangenheit dokumentiert sei (Monahan 1981). Anlass hierfür waren die Auswertungen zweier „natürlicher“ Experimente als Konsequenz zweier Entscheidungen des Obersten Gerichts (Baxter und Dixon), nach denen hunderte von als gefährlich beurteilten Insassen von Hochsicherheitsanstalten, die wegen Gewaltdelikten untergebracht waren, entweder in Freiheit oder in reguläre psychiatrische Einrichtungen entlassen werden mussten (Litwack et al. 2006, 491f.). Die nachfolgende Überprüfung von erneuter Auffälligkeit wegen Gewaltdelikten ergab eine recht geringe Quote von relevanten Vorfällen. Ganz überwiegend hatten sich die Entlassenen in strafrechtlicher Hinsicht unauffällig verhalten. Im Kern ging es um die Frage, ob bei einer extrem geringen „Basisrate“ schwerer Gewalt und einer erheblichen Überschätzung des (Wieder-) Auftretens von Gewalt überhaupt Prognosen durchgeführt werden dürfen (Quinsey 1980).

In der europäischen Kriminologie wurde der Fragestellung der Gefährlichkeit und der Prognose zwar Aufmerksamkeit geschenkt, doch ging die Diskussion kaum über einfache Rückfallprognosen hinaus und auf methodische und theoretische Probleme der Gefährlichkeitsprognose im Engeren ein (Meyer 1965). Ein besonderes Defizit wurde von Frisch (1983) der strafrechtlichen Behandlung der Prognose bescheinigt. Systematische Forschungen mit dem Ziel der Entwicklung von Prognoseinstrumenten und der Analyse der Praxis der Prognosestellung setzten in Deutschland wie auch in anderen europäischen Ländern in den 1990er-Jahren ein. Mit der damals entstehenden Aufmerksamkeit für rückfällige Sexualstraftäter wendeten sich vor allem die forensische Psychologie und Psychiatrie verstärkt der Anpassung und Entwicklung aktuariescher Prognoseinstrumente zu, die zumeist aus den USA oder Kanada übernommen wurden. Erst 2007 kam es in Deutschland zur Veröffentlichung von „Mindestanforderungen an Prognosegutachten“, die von einer Arbeitsgruppe aus Vertretern der forensischen Wissenschaften, der Strafrechtswissenschaften und der Strafjustiz ausgearbeitet wurden (Boetticher et al. 2007). Die Entwicklung von (aktuarischen) Instrumenten, mit denen Straftäter auf ihr Risikopotenzial im Hinblick auf den Rückfall und dessen Schwere eingestuft werden können, wurde in verschiedenen Disziplinen (insbesondere forensische Psychiatrie und Psychologie) sowie Berufsgruppen (Bewährungshilfe, Strafvollzug) wegen des zunehmenden praktischen Bedarfs aufgegriffen und in der jeweiligen Beratungs-, Sachverständigen- und Entscheidungspraxis implementiert (Sjöstedt u. Langström 2001; Nedopil 2005; Dahle 2006; van der Knaap 2009; Belfrage et al. 2012). Neben Anpassungen von in den USA oder Kanada entstandenen Instrumenten kam es nunmehr auch in europäischen Ländern zur Entwicklung von eigenständigen Prognoseinstrumenten (Rossegger et al. 2011). Dabei spielte die Risikoeinschätzung anlässlich von Unterbringungsverfahren (oder Entlassung) bei Geisteskrankheit (bzw. generell

psychiatrisch relevanten Diagnosen, Monahan et al. 2005) eine besondere Rolle. Mit der besonderen rechtspolitischen Aufmerksamkeit für Sexualkriminalität verlagerten sich die Bemühungen nunmehr auf die Entwicklung von spezifischen Prognoseinstrumenten zur sexuellen Gewalt.

1.3 Gefährlichkeit, Prognose und Validität

Die Validität dieser Prognoseinstrumente wird über die Treffsicherheit dargestellt. Zwar ist es richtig, dass die „Prognoseerstellung im forensischen Kontext [...] weniger mit der Frage zu tun hat, ob sie richtig oder falsch im Sinne des späteren Eintritts eines Ereignisses ist“ (Eher et al. 2012), denn die Feststellung eines Risikos von 50% allgemeinen Rückfalls enthält natürlich keine Aussage darüber, ob die dem Risiko zugeordnete Person in die eine oder in die andere Hälfte fällt. Unabhängig davon, wie sich die Person verhält, wird die Prognose (ist sie gemäß den Standards erstellt) dann eben richtig sein und bleiben. Doch ist das Fehlerpotenzial für die Entscheidung über die Anordnung von Maßregeln, ihrer Aussetzung oder die Gestaltung des Vollzugs doch von zentraler Bedeutung, wenn nämlich die Zuordnung zur Risikogruppe mit einer bestimmten Entscheidungsoption verknüpft wird. Für die Entscheider ist ausschlaggebend, mit welcher Häufigkeit bei einer Risikozuordnung die Entscheidung für die Nichtanordnung der Maßregel mit der Konsequenz schwerer Kriminalität verknüpft sein wird und mit welcher Häufigkeit Personen verwahrt werden, obwohl sie in Freiheit keine Straftaten mehr begehen würden. In den grundsätzlichen Aussagen unterscheiden sich die verschiedenen Prognosearten insoweit nicht. Sowohl klinische Prognosen (wie auch immer angereichert durch weitere Instrumente) als auch strikt aktuarische Verfahren kommen immer zu der Aussage, dass einer Person ein hohes, ein mittleres oder ein geringes Rückfallrisiko zuzuordnen ist, oder dass eine Person mit hoher Wahrscheinlichkeit oder mit einer geringeren Wahrscheinlichkeit in der Zukunft Straftaten begehen wird. Die Aussagen können auch auf Begriffe wie Chance, Erwartung oder Gefährlichkeit Bezug nehmen. Die Unterschiede bestehen allein in der Art und Weise der Begründung der Aussage, in der Nachvollziehbarkeit und empirisch belegten Validität von Prädiktoren (des Rückfalls) sowie in der nachvollziehbaren, das heißt überprüfbar Darstellung der Einordnung in numerische Wahrscheinlichkeit und in quantifiziertes Risiko.

Grundsätzlich kann eine Prognose drei Resultate nach sich ziehen: ein richtiges und zwei falsche:

- Das Resultat einer Prognose kann zutreffen.
- Wenn die Prognose mit dem Ereignis nicht übereinstimmt, dann kann der Fehler zwei Formen annehmen.
 - Wurde in der Prognose vorhergesagt, dass ein Ereignis nicht auftreten wird, tritt das Ereignis aber ein, dann wird dieser Fehler als falsch negativer bezeichnet.

- Wurde in der Prognose vorhergesagt, dass ein Ereignis auftreten wird und tritt dieses Ereignis dann nicht ein, dann handelt es sich um einen falsch positiven Fall.

Die Antwort auf die Frage, welcher Fehler reduziert werden soll, hängt nun von einer Reihe von Bedingungen ab, die auf die Art der Gefahr, die politischen Rahmenbedingungen, gesellschaftliche Toleranz, Kosten etc. verweisen. Geht es um die Frage, welche strafrechtliche Sanktion verhängt werden soll oder ob eine Entlassung aus dem Strafvollzug oder einer psychiatrischen Einrichtung erfolgen darf, wird sich eine Konzentration auf die Vermeidung falsch negativer Prognosen ergeben. Dabei wird eine Reduzierung der Zahl falsch negativer zur Zunahme der Zahl falsch positiver führen.

Das Ausmaß an falsch positiven und falsch negativen hängt zuerst von der Basisrate des Ereignisses (und damit theoretische Erklärung) ab, das in bestimmten Gruppen innerhalb eines festgelegten Zeitraums auftreten wird. Die Länge dieses Zeitraums wird ebenfalls Konsequenzen haben. In strafrechtlichen Gesetzen ist lediglich von einer Gefahr oder einer Erwartung die Rede, nicht aber wie lange diese Erwartung oder Gefahr bestehen soll. In diesem Zusammenhang wird immer wieder darauf hingewiesen, dass das Auftreten von schwerer Gewalt auch in stark selektierten Gruppen von wegen Gewaltdelikten (oder gewalttätigen Sexualdelikten) Verurteilten oder Untergebrachten recht selten ist. Insoweit ist auch nachvollziehbar, dass alle statistischen Prognoseinstrumente mit einer hohen Fehlerquote belastet sind. Denn eine niedrige Basisrate führt zu einer Überschätzung des Risikos, vor allem dann, wenn das Ereignis nicht gut erklärt werden kann. Das Problem der systematischen Überschätzung des schweren Rückfalls ist gut bekannt und dokumentiert (Snowden et al. 2007). Dieser Befund deckt sich ohne Weiteres mit Erkenntnissen zur Erklärung von Gewalt, die bislang über recht „weiche“ Theorien nicht hinausgekommen sind (Albrecht 2010). Unabhängig von den statistischen Prognoseinstrumenten ist das Resultat ein hoher Anteil von falsch positiven, der auch deshalb nicht vermieden werden kann, weil die Kriterien nicht theoretisch abgeleitet sind, sondern einfach solche Merkmale abbilden, die mehr oder weniger stark mit dem Wiederauftreten von Kriminalität korrelieren.

Untersuchungen zur Validität von Gefährlichkeitsprognosen sind aus nachvollziehbaren Gründen nur eingeschränkt durchführbar, da die rechtlichen Konsequenzen der Annahme von Gefährlichkeit in der Regel zu langer Freiheitsentziehung führt, die den Nachweis von falsch positiven Fällen nicht erlaubt (Nedopil 2005). Im Wesentlichen bleibt es deshalb bei der Überprüfung von Prognosen bei Entlassung aus dem Strafvollzug oder aus psychiatrischen Einrichtungen und nach einer Begutachtung bzw. Beurteilung auf der Basis klinischer Zugänge oder der Einstufung an Hand von statistischen Prognoseinstrumenten (Quenzer 2010). Auf „natürliche“ Experimente in Form der Entlassung von als gefährlich eingeschätzten Gefangenen wurde weiter oben hin-

gewiesen. Allerdings ergeben sich derartige Untersuchungsansätze selten und wenn, dann für hoch selektierte Gruppen, die lediglich die Widerlegung von Gefährlichkeitsprognosen erlauben, nicht aber die Validierung der Instrumente selbst. Die restriktiven Voraussetzungen für die Anordnung der nachträglichen Sicherungsverwahrung wurden dazu genutzt, die Gefährlichkeitsbeurteilung von Straftätern durch Staatsanwaltschaft und Sachverständige zu überprüfen, die nach Ablehnung der nachträglichen Sicherungsverwahrung entlassen worden waren. Die Analyse des Rückfalls der (zwischen 2004 und 2006) Entlassenen ergab bis Mitte 2008 für 77 Fälle keine erneute Verurteilung in 50 Fällen, 10 Verurteilungen zu Geldstrafe, 5 Verurteilungen zu einer zur Bewährung ausgesetzten Freiheitsstrafe und 12 Verurteilungen zu Freiheitsstrafe ohne Bewährung davon in drei Fällen mit erneuter Anordnung der Sicherungsverwahrung. Darunter befanden sich eine Verurteilung wegen sexueller Nötigung und ein schwerer Missbrauch von Kindern. Hinzu treten eine Verurteilung wegen Körperverletzung und zwei weitere wegen Raubes (5 Fälle aus 77) (Alex u. Feltes 2010, 160f.). In einer weiteren Untersuchung von Entlassenen mit abgelehnter nachträglicher Sicherungsverwahrung, für die eine Gefährlichkeitsdiagnose vorlag, zeigen sich vergleichbare Ergebnisse: Drei Viertel der als gefährlich eingeschätzten Insassen fallen entweder nicht mehr oder nur wegen kleinerer Straftaten auf (Müller et al. 2011). Der Anteil der (Rückfall-)Verteilungen wegen schwerer Gewalttaten entspricht insoweit den Verteilungen in den Baxter und Dixon Fällen; das Ergebnis des „natürlichen“ Experiments im Rahmen der nachträglichen Sicherungsverwahrung stützt die Annahme, dass Gefährlichkeitsbeurteilungen zu einer erheblichen Überschätzung des erneuten Auftretens schwerer Gewalt führen. Vergleiche von als gefährlich beurteilten und deshalb zum Tode verurteilten Gefängnisinsassen mit Verurteilten ohne diese Prognose führen bei langen Beobachtungszeiträumen zu demselben Ausmaß von Gewalt (in Gefängnissen, Edens et al. 2005; Price u. Byrd 2008; Price et al. 2009).

Forschungen zu der Brauchbarkeit von Gefährlichkeitsprognosen (aktuarische Prognosen, klinische Prognosen und strukturierte klinische Prognosen) haben über die letzten Jahrzehnte zu Veränderungen in der Beurteilung der Leistungsfähigkeiten geführt (Murray u. Thomson 2010). Einschätzungen der Treffsicherheit von klinischer und aktuarischer Prognose bzw. von Kombinationen verschiedener Prognosemethoden sind heute optimistischer im Vergleich zu den Beurteilungen aus den 1980er-Jahren (Slobogin 2006, 283ff.). Jedenfalls wird davon ausgegangen, dass sowohl klinische als auch aktuarische Prognosen Ergebnisse liefern, die deutlich über zufällig generierten Resultaten liegen. Jedoch hat sich nichts an der überwiegenden Einschätzung geändert, dass aktuarische Prognosen überlegen sind.

Die Überlegenheit von aktuarischen Prognosen gegenüber klinischer Einschätzung wird seit Jahrzehnten betont und empirisch nachgewiesen. Dies gilt offensichtlich auch für solche Gruppen, für die aktuarische Informationen

nur in begrenztem Umfang vorliegen. So zeigte sich, dass eine aktuarische Beurteilung von Insassinnen einer forensischen Psychiatrie auch dem Zufall hätte überlassen werden können. Die klinische Beurteilung der Gefährlichkeit der Patientinnen führte allerdings zu einer negativen Korrelation, was nichts anderes bedeutet, als dass ungefährliche Patientinnen eher als gefährlich eingeschätzt worden waren und umgekehrt (Harris et al. 2002; vgl. allerdings auch Hood et al. 2002, wo sich Annäherungen ergeben).

Validitätsstudien leiden unter Beschränkungen der Untersuchungsgruppen, die sich weitgehend aus Gefangenen oder in forensischen Einrichtungen Untergebrachten, also hoch selektierten Gruppen, rekrutieren (de Vogel et al. 2004). Sie leiden in besonderem Maße daran, dass für viele Gruppen, insbesondere auch solche, die über psychiatrische Diagnosen definiert sind, die Basisrate von Gewalt, sexueller Gewalt oder Kriminalität allgemein nicht bekannt ist. Der Rückgriff auf allgemeine Rückfallkriminalität, die über Verurteilungen oder Polizeikontakte gemessen wird, verhilft nicht einmal zu einer bescheidenen Annäherung an die Grundverteilungen. Dies gilt insbesondere für den sexuellen Missbrauch, der eine enorme Breite an Handlungen abdeckt, die anhand von Registereintragungen und im Hinblick auf die relative Schwere kaum aufgeschlüsselt werden können. Der Grund liegt wohl einerseits darin, dass sich die forensische Psychologie und Psychiatrie mit den in forensischen Einrichtungen vorhandenen hoch selektierten Gruppen befassen (die in verschiedenen Ländern vollständig unterschiedliche Strukturen aufweisen können), während die allgemeine Kriminalitätsforschung in aller Regel keine Differenzierung entlang psychiatrischer Diagnosen durchführt. Andererseits ist die Basisrate für die schwerste Gewalt, nämlich Tötungsdelikte, so gering (in Deutschland etwa 700 Tötungen bei 82 Millionen Personen im Jahr 2010), dass jede Vorhersage, ist sie nicht beschränkt auf ganz kurze Zeiträume und genau spezifizierte Bedingungen (der Person und der Situation) eigentlich nie besser als der Zufall sein kann.

Vergleichende Untersuchungen zu verschiedenen aktuarischen Prognosen führen im Übrigen nicht zu eindeutigen Ergebnissen (der Überlegenheit des einen oder anderen Instruments). Vielmehr überwiegen Hinweise, dass die Instrumente nur mit äußerster Vorsicht eingesetzt werden sollten und dass weitere Forschung benötigt werde, um die Instrumente für die eigentlich relevanten Untergruppen von Gewalt- oder Sexualstraftäter nutzbar zu machen (Sjöstedt u. Langström 2001, Eher et al. 2012). Auch wenn nur die als Ergebnis der Anwendung aktuarischer Instrumente ermittelten Hoch- und Niedrigrisikogruppen beachtet werden und die besonderer Fehleranfälligkeit ausgesetzte dazwischen liegende Gruppe ausgeblendet wird, ergeben sich Fehleraten einer Größenordnung, die zu einer eher pessimistischen Beurteilung Anlass geben (Menzies et al. 1985, 65). Tiefer gehende und differenzierendere Ergebnisse sind allerdings auch durch zukünftige Forschung schon deshalb nicht zu erwarten, weil gerade die aussagekräftigsten Merkmale (Korrelate

der Gewalt), nämlich frühere Auffälligkeiten bzw. vorangegangene Verurteilungen, in allen Instrumenten gleichermaßen Eingang finden. Gemessen an den Ergebnissen, insbesondere auch in vergleichenden Untersuchungen, geht es offensichtlich auch heute noch nur um die Frage, wer der Stärkste in einer Gruppe von Schwächlingen ist (Freedman 2001). Berücksichtigt man zudem den heutigen Stand der kriminologischen Lebenslaufforschung, dann ist die Zukunft offen. „Shared Beginnings and Divergent Lives“ bedeuten eben, dass dieselben Ausgangsbedingungen zu vollständig unterschiedlichen Lebensverläufen führen können (Laub u. Sampson 2003). Verständlich ist deshalb auch die Empfehlung, Sachverständige sollten sich – wird ihre Meinung in rechtlichen Verfahren eingeholt – auf Hinweise zur Komplexität der Interaktionen zwischen Umwelt und Person beschränken, Prognoseverfahren als vorläufiges Screening Instrument zu nutzen und damit anerkennen sowie kommunizieren, dass in jeder Person ein erhebliches Potenzial der Veränderung angelegt ist (Beecher-Monas u. Garcia-Rill 2006, 340f.). In einer frühen „Amicus Curiae“ Stellungnahme hat sich die Amerikanische Gesellschaft für Psychiatrie in einem der ersten Verfahren vor dem Obersten Gericht zur Frage der Zulässigkeit der Einführung von Sachverständigenaussagen zur Gefährlichkeit gegen die Teilnahme von Psychiatrischen Sachverständigen ausgesprochen. Dabei wurde maßgeblich darauf abgestellt, dass aktuarische Instrumente immer eine entscheidende Grundlage für Gefährlichkeitsprognosen darstellen müssen und dass Psychiater zu Fragen des langfristigen Auftretens von Gewalt auf der Basis aktuarischer Instrumente aus ihrer Ausbildung und vor dem Hintergrund des spezifischen psychiatrischen Wissens keine bedeutsamen interpretativen Beiträge leisten könnten (American Psychiatric Association 1982, 3, vgl. bereits American Psychiatric Association Task Force 1974). Damit wird unterstrichen, dass psychiatrische Krankheiten oder Störungen eben nicht zu den Merkmalen gehören, die theoretisch (nämlich in der Erklärung von Gewalt) oder prognostisch in der Vorhersage von Gewalt von signifikantem Gewicht wären. Vielmehr gehören zu den für eine Prognose brauchbaren Variablen solche, die mit der Psychiatrie und ihren Untersuchungsgegenständen nichts zu tun haben. Zu den aussagekräftigen Merkmalen gehören beispielsweise vorhergehende Auffälligkeiten, insbesondere in Form von Vorstrafen, Alter und Geschlecht, jedenfalls Faktoren, die mit psychiatrisch relevanten Zuständen oder psychiatrischen Krankheiten nicht korrelieren (Monohan 1982, 78; American Psychiatric Association 1982, 5; Steadman et al. 1998). Instrukтив ist insoweit auch eine Stellungnahme der Schweizerischen Akademie der Medizinischen Wissenschaften zur Revision des Schweizerischen Strafgesetzbuches und zur Umsetzung von Art. 123a der schweizerischen Bundesverfassung, der eine lebenslange Verwahrung auch von Straftätern ohne Vorstrafen dann vorsieht, wenn eine hohe Wahrscheinlichkeit der Wiederholung schwerer Straftaten besteht. Im Zentrum der Kritik steht die Feststellung, dass eine solche Prognose von Medizinern und Psychiatern nicht geleistet werden könne, da verlässliche aktuarische Instrumente fehlten und weil ferner psy-

chirurgische Sachverständige ohne Anhaltspunkte für psychische Störungen und insbesondere bei Ersttätern keinerlei Grundlage hätten, eine brauchbare Gefährlichkeitsprognose zu erstellen². Das amerikanische Oberste Gericht hat sich mit der Gefährlichkeitsprognose zwar auseinandergesetzt, die zentralen Probleme aber übergangen. Das Gericht hat argumentiert, die Nichtzulassung von Prognosen käme dem Vorgang gleich, die „Erfindung des Rades rückgängig zu machen“ und bezieht sich dabei selbstverständlich auf die internationale Entwicklung der Sanktionensysteme, die an verschiedenen Stellen des Prozesses eine Prognose fordern. Ferner wurde gesagt, dass Geschworene (und Richter) in der Lage seien, richtige von falschen Prognosen zu unterscheiden (Barefoot v. Estelle, 463 U.S. 880 [1983], vgl. schon Jurek v. Texas, 428 U.S. 262 [1976]). Damit wurden, und dies setzt sich bis heute fort, die an wissenschaftliche Aussagen ansonsten angelegten Maßstäbe für die Gefährlichkeitsprognose nicht herangezogen. Diese Maßstäbe betreffen aussagekräftige Untersuchungen, die die Aussagen stützen, die Prüfung durch die wissenschaftliche Community (peer review), die Einschätzbarkeit von Irrtumswahrscheinlichkeiten und die Akzeptanz der Aussagen in der jeweiligen wissenschaftlichen Disziplin (Daubert Test, Daubert v. Merrell Dow Pharm., 509 U.S. 579 [1993]). Gemessen an solchen Bedingungen ist es um Gefährlichkeitsprognosen schlecht bestellt (Scherr 2003), was freilich nichts daran ändert, dass die amerikanische Rechtsprechung an Gefährlichkeitsprognosen festhält (Slobogin 2006, 293ff.).

Vergleichbare Diskurse zur Gefährlichkeitsprognose sind auch in der obergerichtlichen Rechtsprechung in Deutschland zu erkennen. Der Bundesgerichtshof hat in einer neueren Entscheidung zur Verwendung aktuarischer Prognosemethoden Stellung genommen, die lehrreichen Aufschluss gibt über richterliche Sichtweisen zu Prognosen. Der Bundesgerichtshof hat in dieser Entscheidung ausgeführt, dass sogenannte Strukturprognosen, basierend auf statistischen Modellen, die (über die Zugehörigkeit zu einer wie auch immer definierten Gruppe) Wahrscheinlichkeitsaussagen über das Auftreten von kriminellen Handlungen angeben lassen, nicht ausreichen. Vielmehr sei darüber hinaus eine „fundierte Einzelbetrachtung“ notwendig (BGH, Beschl. v. 30.03.2010 – 3 StR 69/10 = StV 2010, 484; vergleichbar die Entscheidung des Schweizerischen Bundesgerichts zu FOTRES, BGer 6B_772/2007 vom 09.04.2008), die auf einer „sorgfältigen Gesamtwürdigung aller für die Beurteilung der Persönlichkeit des Täters und seiner Taten maßgebenden Umstände“ durch das Gericht vorzunehmen sei. Dabei wird davon ausgegangen, dass statistische Prognoseinstrumente lediglich Anhaltspunkte über die Ausprägung eines strukturellen Grundrisikos liefern könnten. Bei dem Angeklagten handelte es sich um einen nicht vorbestraften, knapp 60-jährigen Mann, der mehrere langjährige Beziehungen zu Frauen hatte, seine durchweg weiblichen Opfer

² www.samw.ch/dms/de/Publikationen/.../Bj_Stlgn_SAMW.pdf