Schriften zur Kriminologie 29

Friedrich Lösel | Lena C. Carl | Eva Link

Evaluation der Behandlung von Sexualtätern im Strafvollzug

Eine umfassende multizentrische Studie

Nomos

Schriften zur Kriminologie

herausgegeben von

Prof. Dr. Katrin Höffler, Universität Leipzig
Prof. Dr. Johannes Kaspar, Universität Augsburg
Prof. Dr. Jörg Kinzig, Eberhard Karls Universität Tübingen
Prof. Dr. Ralf Kölbel, Ludwig-Maximilians-Universität München

Band 29

Friedrich Lösel | Lena C. Carl | Eva Link

Evaluation der Behandlung von Sexualtätern im Strafvollzug

Eine umfassende multizentrische Studie

Nomos

Onlineversion
Nomos eLibrary

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

ISBN 978-3-7560-1152-0 (Print)

ISBN 978-3-7489-1713-7 (ePDF)

1. Auflage 2023

© Nomos Verlagsgesellschaft, Baden-Baden 2023. Gesamtverantwortung für Druck und Herstellung bei der Nomos Verlagsgesellschaft mbH & Co. KG. Alle Rechte, auch die des Nachdrucks von Auszügen, der fotomechanischen Wiedergabe und der Übersetzung, vorbehalten. Gedruckt auf alterungsbeständigem Papier.

Vorwort

In meiner 50jährigen Laufbahn als Wissenschaftler habe ich u.a. in Erlangen, Bielefeld, Nürnberg und Cambridge zirka dreißig aus Drittmitteln finanzierte Forschungsprojekte durchgeführt (vgl. Lösel, 2021). Nur wenige waren so spannend und zugleich schwierig wie jenes über die Wirksamkeit der Behandlung von Sexualstraftätern, über das hier berichtet wird. Bereits die Ausgangslage war ungewöhnlich. Es handelte sich nicht um ein Projekt, für das man unabhängig einen Antrag bei einer Förderinstitution stellt. Vielmehr hatte auf Grund national und international unklarer und manchmal negativer Befunde zur Behandlung von Sexualtätern im Strafvollzug das bayerische Staatsministerium der Justiz angefragt, ob ich eine Evaluation zu diesem Thema in Bayern durchführen wolle bzw. könne. Es war aber auch kein typisches Auftragsprojekt, denn man ließ mir bei der Ausarbeitung des Forschungskonzepts Spielraum und hat sich auch während des Projekts mehr wie ein vertrauensvoller Partner als ein restriktiver Auftraggeber verhalten. Ich hatte mit der Ausarbeitung eines Projektplans etwas gezögert, da ich an der Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) eigentlich im Ruhestand und nur noch unbesoldeter „Senior Professor of Psychology" war/bin, also nicht über die früheren Ressourcen als Lehrstuhlinhaber verfügen konnte. Das Thema war jedoch aus verschiedenen Gründen so reizvoll, dass ich mich auf das Abenteuer dieses Projekts eingelassen habe.

Bereits zum Zeitpunkt unserer Meta-Evaluation der Sozialtherapie in Deutschland (vgl. Lösel et al., 1987) hatte die Straftäterbehandlung in Gefängnissen ein Auf und Ab zwischen unrealistischem Optimismus und Pessimismus erfahren. Wie ich damals geschrieben habe, bleibt jedoch langfristig die Kernfrage, inwieweit man im Gefängnis mit Resozialisierungsmaßnahmen Probanden soweit helfen kann, dass sie anschließend ein Leben in Freiheit ohne Straftaten (vielleicht realistischer ohne schwerwiegende Delikte) führen können. Pauschale Polarisierungen zwischen Resozialisierung und Strafe sind in den vielen Diskussionen zum Thema wenig hilfreich, denn Strafziele wie Schuldausgleich, Sicherheit der Allgemeinheit, negative General- und Spezialprävention (sowie auch Wiedergutmachung) müssen mit der positiven Spezialprävention gemäß StVollzG bzw. seinen länderspezifischen Varianten gleichsam „unter einen Hut" gebracht werden.

Hier hat sich in den letzten 25 Jahren vieles geändert. Denn nachdem die Sozialtherapie als Maßregel der Besserung und Sicherung nicht in Kraft getreten war, führte sie nach der sog. Vollzugslösung ein Schattendasein. Das änderte sich mit der großen Aufmerksamkeit, die Sexualdelikte in der Bevölkerung, den Massenmedien und der Politik erhalten hatten. Dementsprechend verabschiedete der Gesetzgeber in 1998 eine Strafrechtsreform, die einerseits schärfere Sanktionen und andererseits eine Behandlungspflicht in Sozialtherapeutischen Anstalten (SothAs) für Sexualtäter und andere Gewalttäter mit Freiheitsstrafen von mehr als zwei Jahren vorsah. Die Fachwelt äußerte sich damals teilweise skeptisch. Unter anderem stellten sich Fragen zur Behandlungspflicht und auch zu einem realistischen Ausbau der erforderlichen Behandlungskapazitäten. Letzteres ist weitgehend gelungen. Jedoch gab es nach einem systematischen Review damals in Deutschland kaum methodisch gut fundierte Studien, die für einen Erfolg der Sexualstraftäterbehandlung beim sexuellen Rückfall sprachen (vgl. Lösel, 2000). Bei der allgemeinen Rückfälligkeit deuteten sich positive Effekte an, was darauf hinwies, dass die Erfahrungen und Konzepte für die Behandlung der Klientel der Sexualtäter in Gefängnissen hierzulande noch wenig entwickelt waren.

Seither gibt es international und in Deutschland mehr Forschung und Praxiserfahrungen zu diesem Thema. Die Ergebnisse machen aber klare Schlussfolgerungen noch immer schwierig. Internationale Meta-Analysen zeigten zwar im Durchschnitt positive, aber sehr heterogene Ergebnisse (Schmucker & Lösel, 2017). Vereinzelt gab es große Studien, die zu negativen Ergebnissen kamen, d.h. die Unbehandelten wiesen bei der sexuellen Rückfälligkeit geringere Raten auf als die Behandelten (Mews et al., 2017). Zusammen mit Professor Carl Hanson (Kanada) und Statistikern war ich in die Vorab-Begutachtung dieser Studie des britischen Justizministeriums involviert. Wir kamen letztlich zu dem Schluss, dass die Untersuchung zwar manche Mängel hatte (z.B. keine Daten zu Paraphilien), aber methodisch relativ fundiert war. Recherchen britischer Zeitungen und der BBC vor der Veröffentlichung führten zu einem „Skandal", in der die damalige Justizministerin (und später kurzzeitige Premierministerin) Liz Truss bezichtigt wurde, über 100 Millionen Pfund für ein Behandlungsprogramm verschwendet zu haben, das Sexualtäter sogar zu einem größeren Risiko für potentielle Opfer machte. Das betreffende *Sex Offender Treatment Programme (SOTP)* wurde umgehend beendet.

Auch in Deutschland gab es Ergebnisse, die keinen signifikanten Behandlungseffekt von SothAs beim sexuellen Rückfall zeigten (z.B. Buss-

mann & Richter, 2013; Endres, 2015). Bei manchen Risikogruppen deuteten sich sogar negative Tendenzen an. Auch unsere o.g. internationale Meta-Analyse, die nur methodisch aussagekräftige Primärstudien einbezog, erbrachte für die Behandlung im Gefängnis, anders als im ambulanten Bereich, keinen signifikanten (aber keinen negativen) Effekt. Vor diesem Hintergrund mag mein Zögern hinsichtlich der Durchführung der vorliegenden Studie verständlich gewesen sein. Zwar bin ich der empirisch fundierten Überzeugung, dass einzelne Studienergebnisse in der Evaluationsforschung nie die volle „Wahrheit" enthalten und mehr Replikationen auch in der Kriminologie dringend erforderlich sind (vgl. Farrington et al., 2019; Lösel, 2018), denn nicht zuletzt spielen Details der jeweiligen Evaluationsmethodik eine Rolle (vgl. Lösel et al., 2020). Aber ein einzelnes Ergebnis eines großen Projekts kann natürlich starken Einfluss auf die öffentliche Diskussion haben, wie sich bei Mews et al. (2017) gezeigt hat. Ein solches Ergebnis wäre auch in Deutschland schwer zu vermitteln. Natürlich könnte man die Gründe in der ruhigen Atmosphäre universitärer Seminare differenziert analysieren, aber Personen, die sich langjährig in der Behandlung von Sexualstraftätern engagiert haben, wären doch sehr verunsichert und enttäuscht. Auch wären mögliche Reaktionen in sensationslüsternen Medien zu erwarten. Solche hatten wir bereits in unserer BKA-Studie zur Gewalt an Schulen (Lösel & Bliesener, 2003) erlebt, als vor allem gefragt wurde, welche Schulen die meisten Probleme hatten (was von uns unbeantwortet blieb).

Gleichwohl hat mich letztlich meine Überzeugung, dass Evaluationsforschung vor allem eine neutrale empirische Rückmeldung für ein soziales System oder eine Institution sein soll, dazu bewegt, dieses schwierige Projekt durchzuführen. Bestärkt wurde ich darin, dass die vorliegenden Forschungsergebnisse fast durchweg differenzierte Botschaften enthielten, und nicht ein pauschales Urteil wie „Nothing works". Unsere hier berichtete Studie kommt letztlich auch zu einem differenzierten Befund. Etliche Ergebnisse sind für die Sexualtäterbehandlung in bayerischen Gefängnissen durchaus positiv und ermutigend. Bei anderen deutet sich an, dass manche Modifikationen sinnvoll sein könnten. Sie stimmen mit allgemeineren Überlegungen zur schrittweisen Weiterentwicklung der Straftäterbehandlung überein (vgl. Lösel, 2020). Die wichtigsten Befunde der vorliegenden Studie haben wir in der Zusammenfassung dargestellt, doch enthält der Bericht eine Fülle weiterer Ergebnisse, die Interessierte hoffentlich mit Gewinn lesen werden. Trotz entgangener Freizeit im „Ruhestand" lautet mein persönliches Fazit: „Es hat sich gelohnt."

Dieses Fazit war nur möglich, weil zahlreiche Personen und Institutionen das Projekt unterstützt haben. An erster Stelle darf ich meine Mitarbeiterin Dr. Eva Link nennen, die an allen Erhebungen beteiligt war, höchst kompetent Auswertungen durchführte und Entwürfe für diesen Endbericht fertigte. Sie hat dabei zeitweise ihre Promotion zurückgestellt, um unser Projekt erfolgreich abzuschließen. Großen Anteil an unserer Studie hat auch Dr. Lena Carl, die inzwischen auf eine Dauerstelle außerhalb der FAU gewechselt ist, aber an allen Erhebungen und Teilen der Auswertung erfolgreich mitarbeitete. In der Anfangsphase des Projekts wirkte auch Lora Lauchs engagiert mit, die sich dann dem Abschluss ihres Jurastudiums widmen musste. Dr. Martin Schmucker und Dr. Doris Bender vom Institut für Psychologie der FAU leisteten zusätzlich zu ihren sonstigen Aufgaben dem Projekt bei wissenschaftlichen und organisatorischen Fragen Hilfe. Ich danke auch den stundenweise beteiligten Sekretariats- und Hilfskräften sowie der Zentralen Universitätsverwaltung der FAU.

Mein ganz besonderer Dank gilt LRD Dr. Johann Endres und ORRin Dr. Maike Breuer vom Kriminologischen Dienst (KrimD) des bayerischen Justizvollzugs. Von dort erfuhren wir jede benötigte Hilfe und es gab auch keine Probleme durch die etwas ungewöhnliche Konstellation, dass der KrimD einerseits bei uns mitwirkte und anderseits das Ministerium von ihm Kommentare zu unserer Projektarbeit einholte. Es war außerordentlich wertvoll, dass wir die vom KrimD gesammelten Fragebogendaten zur Behandlungspraxis in den bayerischen SothAs auswerten konnten. Diesbezüglich danke ich auch den Fachdiensten, die diese Daten aus ihren Einrichtungen regelmäßig berichtet hatten. Sie ermöglichten einen wertvollen Einblick in frühere Daten, die nicht aktuell erhoben werden konnten. Auch der Kriminologischen Zentralstelle (KrimZ) in Wiesbaden danke ich dafür, dass wir bei Vergleichen mit der früheren Situation der SothAs in Bayern unterstützt wurden. Dank gebührt ebenfalls den Beamten des Bundeszentralregisters, die unseren Antrag auf Datenübermittlung trotz der bei solchen Anfragen üblichen Verzögerung effizient bearbeiteten.

Eine große Hilfe für unser Projekt waren die LeiterInnen der sieben SothAs und die MitarbeiterInnen der Fachdienste und des Allgemeinen Vollzugsdiensts. Trotz immenser Belastungen durch die Corona-Krise waren sie bereit, unsere Befragungen zum Anstalts- und Behandlungsklima zu organisieren und zu bearbeiten. Der Rücklauf war insgesamt sehr erfreulich. Dank gebührt auch jenen Leitern und MitarbeiterInnen der Fachdienste, die sich zusätzlich für intensive Interviews zur Behandlungspraxis bereit erklärten. Ein spezielles "Dankeschön" geht an RDin Dr. Claudia

Groß (Leitung der SothA Erlangen) und Herrn ORR Jörg Fehn (Leitung der SothA Amberg), die uns bei der Entwicklung der Fragebögen und Interview-Leitfäden engagiert berieten. Auch wenn sie dieser Bericht nicht direkt erreichen wird, sei auch allen Insassen der SothAs herzlich gedankt, die sich trotz geringer Belohnung mit einer Tafel Schokolade an unserer Gefangenenbefragung beteiligten. Der Rücklauf ergab fast eine Totalerhebung. Wir planen eine Kurzinformation über unsere Ergebnisse für interessierte Gefangene.

Schließlich darf ich dem bayerischen Staatsministerium der Justiz und den im Lauf der Zeit mit unserem Projekt befassten Personen danken, namentlich Herrn MDg Holzner, Ltd MR Krä, MR Dr. Koch-Schulte, MR Dr. Hagen, RD Dr. Witzigmann, MRin Mühlbauer und posthum MR Thomas. Auch wenn Videokonferenzen in Corona-Zeiten bei mir nicht immer perfekt funktionierten, zeigte man im BayStMJ stets Verständnis und Geduld mit unserer Projektarbeit. Ich hoffe, dass man auch dort nach Lektüre des vorliegenden Berichts zu der Auffassung gelangt: „Es hat sich gelohnt."

Erlangen-Nürnberg, im Mai 2023 Friedrich Lösel

Literatur:

Bussmann, K.-D., & Richter, K. (2013). *Kriminologische Evaluation der Sozialtherapeutischen Anstalt Halle (Saale). Abschlussbericht an die DFG*. Halle (Saale): Martin-Luther Universität.

Endres, J. (2015). *Sexualstraftäterbehandlung und Rückfall – Forschungsüberblick und Ergebnisse der bayerischen Sexualtäter-Untersuchung.* Unveröffentlichter Bericht. Erlangen: Kriminologischer Dienst des bayerischen Justizvollzugs.

Farrington, D.P., Lösel, F., Boruch, R.F., Gottfredson, D.C., Mazerolle, L., Sherman, L.W., & Weisburd, D. (2019). Advancing knowledge about replication in criminology. *Journal of Experimental Criminology, 15, 373–396.*

Lösel, F. (2000). The efficacy of sexual offender treatment: A review of German and international evaluations. In P.J. van Koppen & N.H.M. Roos (Eds.), *Rationality, information and progress in psychology and law* (pp. 145–180). Maastricht, NL: Metajuridica.

Lösel, F. (2018). Evidence comes by replication, but needs differentiation: The reproducibility issue in science and its relevance for criminology. *Journal of Experimental Criminology, 14*, 257–278.

Lösel, F. (2020). Entwicklungspfade der Straftäterbehandlung: Skizzierte Wege und Evaluation der Zielerreichung. *Forensische Psychologie, Psychiatrie, Kriminologie, 14,* 35–49.

Lösel, F. (2021). From childhood in a ruined German city to research on crime and violence. In R. Tremblay (Ed.), *The science of violent behavior development and its prevention: Contributions of the Second World War generation* (pp. 271–298). Cambridge, UK: Cambridge University Press.

Lösel, F., & Bliesener, T. (2003). *Aggression und Delinquenz unter Jugendlichen: Untersuchungen von kognitiven und sozialen Bedingungen.* Neuwied: Luchterhand.

Lösel, F., Link, E., Schmucker, M., Bender, D., Breuer, M., Carl, L., Endres, J., & Lauchs, L. (2020). On the effectiveness of sexual offender treatment in prison: A comparison of two different evaluation designs. *Sexual Abuse: A Journal of Research and Treatment, 32*, 452–475.

Lösel, F., Köferl, P., & Weber, F. (1987) *Meta-Evaluation der Sozialtherapie*. Stuttgart: Enke.

Mews, A., Di Bella, L., & Purver, M. (2017). *Impact evaluation of the prison-based core sex offender treatment programme.* London: UK Ministry of Justice, Analytical Series.

Schmucker, M., & Lösel, F. (2017). Sexual offender treatment for reducing recidivism among convicted sex offenders: a systematic review and meta-analysis. *Campbell Systematic Reviews*. Oslo, NOR: The Campbell Collaboration

Inhaltsverzeichnis

1 Zusammenfassung / Executive Summary 15

2 Forschungsstand der Behandlung von Sexualstraftätern und Themen unseres Projekts 31

2.1 Deliktarten, Tätertypen und Behandlungsmethoden bei Sexualdelinquenz 31

2.2 Evaluationsstudien zur Behandlung von Sexualtätern 39

2.3 Organisation und Themen unseres Projekts 52

3 Datenerhebungen 59

3.1 Erhebung struktureller und therapeutischer Rahmenbedingungen 59

3.2 Erhebung von Rückfalldaten und Bestimmung der Untersuchungsstichprobe für die Rückfallanalysen 60

3.3 Kodierung der Deliktschwere 65

3.4 Erhebung des Abteilungs- und Behandlungsklimas 67

3.4.1 Befragung der Bediensteten 67

3.4.2 Befragung der Gefangenen 71

3.5 Interview-Erhebung zu Therapieinhalten und -abläufen 74

4 Datenanalyse und Ergebnisse 77

4.1 Strukturelle und therapeutische Rahmenbedingungen der Therapie 77

4.2 Deskriptive Analysen der Untersuchungsstichprobe für die Rückfallanalysen 85

4.2.1 Therapieteilnahme 85

4.2.2 Biografische Merkmale 87
4.2.3 Psychologische und psychiatrische Daten 88
4.2.4 Kriminologische Daten 90
4.2.5 Merkmale der Sexualdelikte 91
4.2.6 Vollzugs- und Behandlungsverlauf 94
4.2.7 Merkmale des sozialen Empfangsraums 95
4.2.8 Rückfälligkeit 96

4.3 Analysen der Behandlungswirksamkeit hinsichtlich des dichotomen Rückfalls 97
4.3.1 Beschreibung der Analysemethode 98
4.3.2 Vergleich von behandelten und unbehandelten Sexualstraftätern 101
4.3.3 Untersuchung von Kohorteneffekten 106
4.3.4 Vergleich von SothA-Teilnehmern und unbehandelten Sexualstraftätern 109
4.3.5 Vergleich der im Normalvollzug behandelten und unbehandelten Sexualstraftäter 112
4.3.6 Behandlungseffekte bei verschiedenen Tätertypen 115

4.4 Untersuchung differenzierter Rückfallkriterien 122
4.4.1 Zeitliche Dauer bis zum Rückfall 122
4.4.2 Untersuchung der Rückfallschwere 126
4.4.2.1 Pilotstudie zur Untersuchung der Schwere der Sexualdelikte 127
4.4.2.2 Schwere-Analysen zum allgemeinen Rückfall in der erweiterten Stichprobe 131
4.4.2.3 Schwere-Analysen zur Sexualdelinquenz in der erweiterten Stichprobe 134

4.5 Unterschiede zwischen den sozialtherapeutischen Einrichtungen 135
4.5.1 Klientel und Rückfallraten 135
4.5.2 Behandlungsklima 139
4.5.2.1 Analysen zur Befragung der Bediensteten 139
4.5.2.2 Analysen zur Befragung der Gefangenen 152
4.5.2.3 Zusammenhänge zwischen Abteilungsklima, Rückfälligkeit und Therapieabbruch 165
4.5.3 Therapieabläufe und -inhalte 168
4.5.3.1 Eingangsphase/ Therapiebeginn 168
4.5.3.2 Therapie 175
4.5.3.3 Entlassung 187
4.5.3.4 Schlussbemerkungen 192
4.6 Weitere Analysen 192
4.6.1 Analysen zu Therapieverweigerung und -abbruch 192
4.6.2 Analysen zu Therapieauflagen und Nachsorge 198
4.6.3 Analysen zu Vollzugslockerungen 201
4.6.4 Problemstellungen und Motivationsmaßnahmen in der Behandlung von Sexualstraftätern 207
4.6.5 Auswirkungen der Covid-19-Pandemie 212

5 Diskussion 219
5.1 Allgemeine Charakteristika der Studie 219
5.2 Ergebnisse der Analysen zur Rückfälligkeit 222
5.3 Unterschiede zwischen den Anstalten und Organisationsklima 228
5.4 Abläufe und Inhalte der Behandlung 234
5.5 Probleme in der Behandlung 237
5.6 Auswirkungen der Covid-19-Pandemie 239
5.7 Limitationen 240
5.8 Schlussbemerkung 242

Anhang 255
Anhang 1 – Fragebogen zur Erhebung der strukturellen und therapeutischen Rahmenbedingungen 255

Anhang 2 – Erhebungsbogen für Sexualstraftäter 270
Anhang 3 – Kodierschema zur Kodierung der BZR-Auszüge 281
Anhang 4 – Items/Skalen der Klima-Erhebung bei den Bediensteten 285
Anhang 5 – Items/Skalen der Klima-Erhebung bei den Gefangenen 287
Anhang 6 – Interviewleitfaden zur Erhebung von Therapiemerkmalen 290
Anhang 7 – Standardisierte Differenzen in der ursprünglichen und der PS-gewichteten Stichprobe bei behandelten und unbehandelten Probanden 298
Anhang 8 – Standardisierte Differenzen in der ursprünglichen und der PS-gewichteten Stichprobe bei SothA-Teilnehmern und unbehandelten Probanden 302
Anhang 9 – Standardisierte Differenzen in der ursprünglichen und der PS-gewichteten Stichprobe bei im Normalvollzug behandelten (ET/GT) und unbehandelten Probanden 306
Anhang 10 – Standardisierte Differenzen in der Teilstichprobe der behandelten und unbehandelten Kindesmissbraucher 311
Anhang 11 – Standardisierte Differenzen in der Teilstichprobe der behandelten und unbehandelten Vergewaltiger 315
Anhang 12 – Interne Konsistenzen (Cronbachs α) der Skalen der Klima-Erhebung 319

1 Zusammenfassung / Executive Summary[1]

Zusammenfassung

Die Wirksamkeit der Behandlung von Sexualstraftätern in Gefängnissen wird international kontrovers diskutiert. Systematische Auswertungen und Meta-Analysen von Evaluationsstudien zeigen unterschiedliche und im Durchschnitt eher mäßig positive Effekte. Aus einigen Studien, z.B. in Großbritannien und Deutschland, werden auch tendenziell negative Ergebnisse berichtet, d.h. die Behandlungsgruppe (BG) hatte bei sexuellen Delikten sogar eine etwas höhere Rückfallrate als die unbehandelte Kontrollgruppe (KG). Es ist erfreulich, dass international und auch in Deutschland sexuelle Rückfalldelikte insgesamt zurückgegangen sind. Dies bedeutet zugleich niedrige Basisraten, bei denen statistisch signifikante Unterschiede zwischen BG und KG schwer nachweisbar sind. Insgesamt ist in der Sexualtäterbehandlung eine differenzierte Sicht und Evaluationsmethodik erforderlich.

Vor diesem Hintergrund hat das bayerische Staatsministerium der Justiz das hier berichtete Forschungsprojekt angeregt und gefördert. Es handelt sich nach unserer Kenntnis um das bislang umfangreichste *abgeschlossene* Projekt zur Evaluation der Sexualtäterbehandlung in sozialtherapeutischen Anstalten bzw. Abteilungen (SothAs) des Justizvollzugs. In den sieben SothAs wurden nicht nur die (offizielle) Rückfälligkeit untersucht, sondern z.B. deren Art, Schweregrad und zeitlicher Verlauf. Darüber hinaus befasste sich die Studie mit den Behandlungsinhalten, Prozessen der Diagnostik und Probandenauswahl, dem Sozialklima in den Abteilungen aus der Sicht des Personals und der Gefangenen, Problemen der Therapie, der Gewährung von Lockerungen, der therapeutischen Nachsorge in den Ambulanzen, spezifischen Einflüssen durch die Corona-Pandemie und anderen Aspekten.

1 Wir verwenden in diesem Bericht zur sprachlichen Vereinfachung den weithin gebrauchten Begriff „Sexualstraftäter". Dabei sollte aber berücksichtigt werden, dass es sich um Personen mit vielfältigen anderen Charakteristika handelt, die sich nicht auf die sexuelle Anlasstat reduzieren lassen. Genauer, aber umständlich, wäre die Bezeichnung „Personen, die eine sexuelle Straftat gemäß StGB begangen haben."

Unser Projekt erweiterte die vom Kriminologischen Dienst des bayerischen Justizvollzugs durchgeführte Studie, sodass nun aus 1.787 Auszügen des Bundeszentralregisters valide Rückfalldaten für 1.245 Sexualtäter aus den Jahren 2004 bis 2015 vorlagen. Da sich behandelte und unbehandelte Probanden in verschiedener Hinsicht unterschieden, wurde ein Propensity Score Matching (PSM) durchgeführt, das zahlreiche für die Behandlung und den Rückfall relevante Merkmale einbezog. Dadurch konnten Selektionseffekte vermieden und die BG und KG als äquivalent angesehen werden. Im Sinne einer konservativen Abschätzung von Therapieeffekten folgten wir einem Intent-to-Treat-Ansatz, wodurch Probanden, die die Behandlung nach der Probephase abgebrochen hatten, der BG zurechnet wurden. Die Stichprobe für die Rückfallanalysen umfasste 710 Probanden in der BG und 535 Probanden in der KG. Die durchschnittliche Bewährungszeit in Freiheit nach der Entlassung (Follow-up) betrug M = 9,3 Jahre (SD = 3,5 Jahre). Die PSM-Analysen zeigten für den Rückfall (RF) insgesamt einen statistisch signifikanten Behandlungseffekt (BG = 35,9 % vs. KG = 41,4 %). Tendenzen in die erwünschte Richtung bestanden beim RF mit einem Sexualdelikt (6,0 % vs. 8,0 %), mit einem Gewaltdelikt (13,6 % vs. 14,3 %) oder einem „schweren" Delikt mit erneuter Freiheitsstrafe von mehr als zwei Jahren bzw. Maßregel-Unterbringung (7,5 % vs. 9,1 %). Ein Vergleich der früheren Entlassungskohorten (2004–2008) aus der Untersuchung des KrimD mit den Kohorten aus unserer erweiterten Stichprobe (2009–2015) zeigte in allen vier genannten Kriterien niedrigere Werte bei den später Behandelten. Der sexuelle RF war dabei mit 4,1 % noch niedriger als früher. Die Tendenz zu niedrigeren RF-Raten bestand auch dann, wenn der Rückfallzeitraum jeweils auf drei Jahre vereinheitlicht wurde. Separate Rückfall-Analysen für die neueren Kohorten mittels PSM kamen in der Tendenz zu ähnlichen Ergebnissen wie die Analysen der früheren Kohorten. Für den allgemeinen Rückfall zeigte sich ein statistisch signifikanter positiver Behandlungseffekt. In der Gesamtschau weisen die Ergebnisse der Kohorten-Vergleiche auf eine im Lauf der Zeit verbesserte Behandlungspraxis hin. Die Zusammensetzung der Gruppen hatte sich kaum verändert und war somit keine hinreichende Erklärung der Kohorteneffekte. Man kann aber einen positiven Einfluss günstigerer allgemeiner Rahmenbedingungen für die Resozialisierung nicht ausschließen, was in unserer Studie nicht kontrollierbar war.

Um die Aussagekraft des dichotomen RFs zu differenzieren, haben wir einen von der Sanktionsbiographie unabhängigen Schwereindex (Harm Index) entwickelt, der auf dem mittleren gesetzlichen Strafrahmen basiert.

In der BG zeigte sich eine Tendenz des stärkeren Rückgangs in der Deliktschwere als in der KG. Wir untersuchten auch den Zeitverlauf bis zum ersten Rückfall, da ein späterer Rückfall im Sinne der abnehmenden Alterskurve der Kriminalität zu selteneren Straftaten in der Zukunft führen kann. Die Analyse der Zeit bis zum ersten Rückfall nach der Entlassung ergab, dass die BG signifikant später allgemein rückfällig wurde. Ähnlich war die Tendenz beim Gewalt-RF, während sie beim sexuellen RF leicht umgekehrt war, was vermutlich mit der höheren Anzahl an Missbrauchstätern und Paraphilien in der BG zusammenhängen dürfte.

Obwohl die Behandlung in einer SothA die gesetzliche Norm ist, gab es in der BG auch eine Teilstichprobe von 261 Probanden, die nicht in einer SothA behandelt worden waren, sondern Einzel- oder Gruppentherapie im Normalvollzug erhalten hatten (z.B. weil sie zur Verlegung in eine SothA nicht bereit waren oder andere Hinderungsgründe bestanden). Auch hier zeigte eine PSM-Analyse, dass in den vier RF-Kriterien die Behandelten jeweils seltener rückfällig wurden. Auf Grund der geringeren Fallzahlen waren die Unterschiede aber nicht signifikant. Das Muster der Ergebnisse war jedoch einheitlicher und positiver als in einer PSM-Analyse, die nur die SothA-Probanden einbezog. Hierbei schnitt beim Gewalt-RF und schweren Rückfall die KG etwas besser ab. Diese Befunde waren zwar nicht signifikant, sind aber ein Anzeichen dafür, dass die gesetzlich vorgesehene Behandlungspflicht in einer SothA eventuell nicht für alle Fälle optimal ist. Für eine genauere Interpretation wären Details der im Normalvollzug durchgeführten Einzel- und Gruppentherapien nötig, die nicht vorlagen.

Für die RF-Raten nach Behandlung in den SothAs spielten die Abbrüche in der Intent-to-treat-Analyse teilweise eine negative Rolle. Probanden, die die Behandlung regulär beendeten, schnitten tendenziell besser ab als die Gesamtgruppe und insbesondere als die Subgruppe der Abbrecher. Dies traf vor allem und signifikant für den allgemeinen RF zu (35 % vs. 52 % RF). Allerdings zeigen unsere Analysen, dass nicht nur der Abbruch als solcher, sondern auch damit zusammenhängende Probandenmerkmale und andere Faktoren zu berücksichtigen sind.

Was die Indexdelikte der Probanden betrifft, hatten wegen Vergewaltigung verurteilte Probanden deutlich höhere Raten beim allgemeinen RF, Gewalt-RF und schweren RF, aber etwas niedrigere Raten beim sexuellen RF. Diese Befunde ergaben sich sowohl für Behandelte als auch Unbehandelte. Eine kleinere Teilstichprobe hatte sowohl Kindesmissbrauch als auch Vergewaltigung Erwachsener begangen („gemischte" Täter). Die drei Gruppen unterschieden sich in etlichen Merkmalen u.a. hinsichtlich der

Verübung sonstiger Gewaltdelikte, Art und Zahl der Opfer und früheren Sexualdelikten. Das Profil der „gemischten" Täter ähnelte dem der Vergewaltiger und auch die RF-Raten waren jeweils ähnlich. Separate Rückfallanalysen für Missbrauchstäter und Vergewaltiger (mittels PSM) zeigten in der Tendenz für letztere etwas positivere Ergebnisse und durchweg geringere Rückfallraten von bis zu vier Prozentpunkten bei den Behandelten. Bei Missbrauchstätern waren der allgemeine und sexuelle RF in der BG etwas höher als in der KG. Für beide Tätertypen waren die Unterschiede in den Rückfallraten bei Behandelten und Unbehandelten allerdings nicht signifikant.

Wesentliche Aspekte für die Resozialisierung sind sukzessive Lockerungen und eine effektive Nachsorge. Hier zeigte sich, dass entsprechend dem Konzept der Sozialtherapie zirka zwei Drittel der Behandelten Lockerungen erhalten hatten, bei den Unbehandelten war dies nur gut ein Drittel. Auch bei Kontrolle diverser Risikovariablen ging die Gewährung von Lockerungen mit geringeren RF-Raten einher. Durch den Ausbau der ambulanten Therapieeinrichtungen sind diese ein wesentlicher Bestandteil der Resozialisierung geworden, sodass über drei Viertel der Sexualtäter in unserer Studie eine Therapieweisung erhalten hatten. Dies war u.a. etwas häufiger bei Tätern mit kindlichen Opfern und einem niedrigeren (nicht höherem!) statischen RF-Risiko (Static-99) der Fall. Hinsichtlich der Rückfälligkeit bestand jedoch kein „Booster-Effekt", was auf eine gewisse therapeutische „Sättigung" hinweist. Dagegen zeigten sich positive Effekte der ambulanten Nachbehandlung bei den im Gefängnis unbehandelten Probanden. Als Ansatzpunkte für eine Verbesserung der Nachsorge wurden von den Befragten der Fachdienste eine strukturiertere Übergabe an die Fachambulanzen, eine flexiblere Anpassung der Nachsorge an individuelle Behandlungsbedürfnisse sowie ein räumlicher Ausbau des Versorgungsnetzes genannt.

Bereits in einer Studie an den Entlassungskohorten bis 2008 hatten wir Unterschiede in der Klientel und in den RF-Raten der sieben SothA-Einrichtungen festgestellt. Unter Kontrolle von Probandenmerkmalen ergab sich dabei, dass die RF-Unterschiede überwiegend mit dem Risikograd der Insassengruppen zusammenhingen. In der erweiterten Stichprobe widmeten wir uns nun detaillierter den Unterschieden zwischen den Einrichtungen. Dazu wurden Befragungen der Bediensteten und Gefangenen zum Sozial- und Behandlungsklima in den SothAs durchgeführt. Die international erprobten Klimaskalen des EssenCES, SQL (Bedienstete) und MQPL (Gefangene) betrafen u.a. die Einstellung zu den Gefangenen, das Sicher-

heitserleben, die Unterstützung der Probanden durch die Bediensteten, die Beziehungen zu KollegInnen, den Zusammenhalt der Gefangenen und die Motivation (Personal) sowie das Wohlbefinden, die persönliche Entwicklung, den erlebten Stress und die Kontakte zur Familie (Gefangene). Die anonymen Befragungen erbrachten einen sehr befriedigenden Rücklauf (Personal 73 %, Gefangene 91 %). In den meisten Dimensionen lagen die durchschnittlichen Angaben im mittleren oder leicht positiven Bereich der Skalenbreite und hatten eine plausible Streuung. Das heißt, die Antworten erfolgten nicht pauschal positiv oder negativ. Beim Personal war die Einschätzung der Gefangenen im Fachdienst etwas positiver als beim Allgemeinen Vollzugsdienst. Der Gesamtwert im EssenCES und positive Werte hinsichtlich der Unterstützung der Gefangenen sowie der Sicherheit und Kontrolle in der Einrichtung hingen u.a. signifikant mit der Einschätzung der Wirksamkeit der Behandlung zusammen. Noch ausgeprägter (bis zu r = 0,62) waren die Korrelationen zwischen der globalen Bewertung der Arbeit in der SothA und dem SQL-Gesamtwert, der Einstellung zur Leitung, der eingeschätzten Sicherheit und des Zusammenhalts der Gefangenen.

Zwischen den Angaben des Personals und der Gefangenen bestanden in den einzelnen Dimensionen sowohl Ähnlichkeiten als auch Unterschiede. Am deutlichsten waren die Unterschiede hinsichtlich der Unterstützung durch Bedienstete in drei Einrichtungen, während in anderen in beiden Gruppen ähnliche Werte vorlagen. Es traf auch nicht zu, dass einzelne SothAs in allen Skalen negativere Klimawerte aufwiesen, sondern das Bild war differenziert. Mit verschiedenen Skalen korrelierten u.a. das Alter der Gefangenen und die Art der SothA, in der sie untergebracht waren (nur Sexualtäter oder zusammen mit sonstigen Gewalttätern). Insbesondere die Erfahrung von Drogen und Ausbeutung, die erlebte Sicherheit und der Zusammenhalt unter den Gefangenen waren in Einrichtungen mit gemischter Klientel negativer. In diesem Ergebnis deuten sich jene Stigmatisierungsprobleme an, die vor allem Gefangene mit Kindesmissbrauch im Normalvollzug haben.

Auch in unserer erweiterten Stichprobe gab es hochsignifikante Unterschiede in den RF-Raten zwischen den sieben Abteilungen, und zwar beim allgemeinen RF und Gewalt-RF. Hier waren in manchen Einrichtungen die RF-Raten doppelt bis dreimal so hoch wie in anderen. Beim sexuellen RF wurden auch wegen der geringen Basisraten die Unterschiede nicht signifikant. Erwartungsgemäß war der durchschnittliche Risikograd der jeweiligen Klientel ein zentraler Faktor für die RF-Unterschiede zwischen den SothAs. Es zeigte sich aber auch, dass verschiedene Klimaskalen mit

dem RF zusammenhingen. So ging beim allgemeinen RF eine positivere Wahrnehmung des Personals mit weniger Rückfällen der jeweiligen Einrichtung einher. Mit niedrigeren Raten beim Gewalt-RF korrelierten positivere Beurteilungen der Gefangenen hinsichtlich der Unterstützung durch Bedienstete, eine positivere Wahrnehmung des Personals und ein laut Bediensteten stärkerer Zusammenhalt der Gefangenen. Diese Zusammenhänge bestanden auch bei Kontrolle des RF-Risikos.

In ausführlichen Interviews mit MitarbeiterInnen der Fachdienste wurden Details der Behandlung untersucht. Zwar bilden die gesetzlichen Regelungen einen einheitlichen Rahmen für alle SothAs, doch gab es in der konkreten Ausgestaltung neben den Ähnlichkeiten zwischen den sieben Einrichtungen auch nicht unerhebliche Unterschiede. Diese zeigten sich z.B. bei der Auswahldiagnostik, der Häufigkeit von Einzel- und Gruppentherapie-Sitzungen, dem Einsatz strukturierter Programme, der Gewährung von Lockerungen, den Kontakten mit Angehörigen und der Vorbereitung der ambulanten Nachbetreuung. Zwar orientieren sich die Abteilungen überwiegend auch am international bewährten Risk-Need-Responsivity-Prinzip als Grundlage der Behandlung, dessen Ausgestaltung ist aber recht unterschiedlich. Dies beruht offenbar auf den jeweiligen praktischen Erfahrungen und spezifischen lokalen Präfenzen. Hierzu wäre ein Austausch zu einzelnen Themen und eine praxisnahe empirische Evaluation sinnvoll. Beispielsweise stellt sich die Frage, ob Leugnen ein Ausschlussgrund für die Therapieteilnahme sein sollte, was aufgrund der Deliktarbeit als integraler Bestandteil der meisten Therapiekonzepte überwiegend der Fall ist. Internationale und deutsche Studien zeigen, dass das Leugnen jedoch kaum mit dem RF zusammenhängt.

In den Interviews sowie der Fragebogen-Erhebung wurden auch ausführlich Probleme der Behandlung angesprochen. Diese wurden primär bei den Gefangenen wahrgenommen, z.B. mangelnde Veränderungsmotivation, zu wenig Mitarbeit in der Therapie, disziplinarische Verfehlungen, Weitergabe vertraulicher Informationen, keine ernsthafte Auseinandersetzung mit dem eigenen Sexualdelikt, Leugnen von Aspekten der Tat, Konflikte zwischen Gefangenen, geringe Fähigkeit zur Introspektion und Empathie, kognitive Einschränkungen, begrenzte Sprachkenntnisse, Persönlichkeitsstörungen sowie enttäuschte Erwartungen hinsichtlich der Gewährung von Lockerungen oder vorzeitiger Entlassung.

In unserem Projekt wurden auch zahlreiche weitere Daten zu den SothAs, zum Personal und zu den Gefangenen erhoben, die in einer kurzen Zusammenfassung nicht adäquat berichtet werden können. An einem

bestimmten Thema näher Interessierte seien auf das detaillierte Inhaltsverzeichnis verwiesen. So zeigte sich z.B. bei den Merkmalen der Gefangenen u.a., dass in der Probandenauswahl ein etwas stärkerer Fokus auf Missbrauchstätern liegt. Beim Personal legen die Daten u.a. nahe, dass die große Mehrheit mit ihrer Arbeit sehr zufrieden ist. Sowohl beim Fachdienst als auch beim AVD bestand aber der deutliche Eindruck, dass die Arbeit in der Sozialtherapie bei den Bediensteten im Normalvollzug zu wenig anerkannt wird. Hier wäre eventuell ein realistischer Informationsaustausch angezeigt. Weitere Details in unserem Bericht beziehen sich auf die objektiven Rahmendaten der SothAs in Bayern. Darüber hinaus befassten wir uns auch mit den negativen Einflüssen durch „Corona". Hier erwiesen sich u.a. die Einschränkungen bei Gruppentherapien, die soziale Isolation und mangelnde Außenkontakte als besonders ungünstig.

Trotz unserer ziemlich umfassenden Erhebungen und Datenanalysen bestehen wie in jeder Studie natürlich Einschränkungen. Dazu gehört zum einen das Problem, dass wir nicht wie bei einem randomisierten Kontrollgruppen-Design bereits vor dem Beginn der Behandlung ansetzen und die Personen in der BG und KG zufällig zuweisen konnten. Dies war und ist aus rechtlichen und praktischen Gründen nicht möglich und auch international ein typisches Problem. Durch unsere PSM-Analyse konnten wir aber ein stichhaltiges quasi-experimentelles Design realisieren, das valide Daten zur Rückfälligkeit geliefert hat.

Eine zweite Limitation besteht darin, dass wir wiederholt statistisch nicht signifikante Tendenzen und Trends berichtet und (vorsichtig) interpretiert haben. Dies war zum einen dadurch begründet, dass wir konservativ statistisch zweiseitig getestet und im Intent-to-Treat-Ansatz Behandlungsabbrüche der BG zugerechnet haben. Zum andern sind allzu ritualisierte statistische Signifikanztestungen seit langem auch unter Statistikern sehr umstritten. Bloße Signifikanzangaben können je nach Thematik, Kontext und Resultaten auch irreführen. Gerade in dem noch nicht konsolidierten Feld der Sexualtäterbehandlung war es deshalb wichtig, auch absolute Differenzen zwischen BG und KG sowie konsistente Muster von Ergebnissen zu berücksichtigen. Dies ist zudem dadurch gerechtfertigt, dass wir eine große Stichprobe untersuchten und mehr statistische „Power" nicht möglich gewesen war, da wir für die untersuchten Zeiträume ohnedies eine Totalerhebung durchführten.

Eine dritte Einschränkung in unserer Studie besteht darin, dass die Inhalte und Probleme der Behandlung, das Institutionsklima und andere Daten in den letzten Jahren erhoben wurden, also nicht notwendig für

die Situation vor fünf Jahren zutreffen müssen. Dies war methodisch unvermeidlich und ist ein häufiges Problem von Quasi-Experimenten mit einem langen Follow-up. Wir haben aber aus einem Vergleich mit früheren Erhebungen zu den SothAs Anhaltspunkte dafür, dass auf struktureller Ebene und hinsichtlich von Merkmalen des Personals und der Klientel keine fundamentalen Veränderungen eingetreten sind. Dass sich in den neueren Kohorten die Behandlung als etwas wirksamer als früher erwies, ist als plausibler Lern- und Entwicklungsprozess zu sehen und unterstreicht die Validität unserer Studie.

Eine weitere Einschränkung besteht darin, dass wir uns in unseren Rückfallanalysen auf die offiziellen Daten aus dem BZR beschränken mussten. Nachbefragungen lange nach der Entlassung waren nicht möglich. Zudem sind Täterbefragungen im Dunkelfeld gerade bei der Sexualdelinquenz problematisch. Forschungspraktisch war es auch nicht möglich, neben der Rückfalldelinquenz andere Merkmale der Lebensführung nach der Entlassung zu erfassen, die aber in der Forschung mit dem juristisch zentralen Rückfallkriterium zusammenhängen. Durch unsere differenzierte Erfassung des Rückfalls hinsichtlich Schweregrad, zeitlichen und anderen Merkmalen geht unsere Studie über einfache dichotome Rückfallkriterien hinaus. Bei den Effektstärken zum offiziellen Rückfall muss auch berücksichtigt werden, dass es sich um sog. „single act criteria" handelt, die nur begrenzt etwas über Persönlichkeitsentwicklungen aussagen können.

Wie es bei einer praxisorientierten fairen Evaluation angemessen ist, haben wir unseren Forschungsbericht an die teilnehmenden Institutionen vor der hier vorliegenden Veröffentlichung rückgemeldet. In einem Workshop wurden Argumente über die Ergebnisse und ihre möglichen Gründe aus der Sicht der Praxis ausgetauscht. Schwerpunkte der Diskussion war insbesondere der Behandlungsabbruch (seine Anlässe und Folgen) sowie die Variation der Behandlungsinhalte in den verschiedenen Einrichtungen. Ein in unseren Interviews noch nicht rückgemeldeter Aspekt betraf die Behandlung von Sicherungsverwahrten, die auf Grund der einschlägigen gesetzlichen Regelungen in der Zukunft quantitativ und qualitativ ein größeres Problem für die SothAs darstellen dürften. Die hilfreichen Argumente in den Diskussionen mit den Praxisvertretern wurden von uns noch in die vorliegende Buchpublikation eingearbeitet. Nach unserem Eindruck hat man in der Praxis unsere Studie konstruktiv aufgenommen und erste Schritte zur Umsetzung mancher Befunde initiiert. Insgesamt wird die Aussagekraft unseres Projekts auch dadurch bestätigt, dass einige Ergebnisse

bereits in fünf Zeitschriften mit striktem Peer-Review veröffentlicht werden konnten.

Summary

International discussions of the effectiveness of treatment of sexual offenders in prisons are still controversial. Systematic reviews and meta-analyses of evaluation studies revealed on average moderately positive but very heterogeneous effects. A few studies, e.g. in Great Britain and Germany, even reported slightly negative effects; i.e. the treatment group (TG) had somewhat higher recidivism rates in sexual reoffending than the untreated control group (CG). It is good that over the last decades the rate of sexual reoffending declined internationally and in Germany. This led to lower base rates for which it is difficult to prove statistically significant differences between TG and CG. Overall, differentiated views and evaluation methods are necessary.

Against this background, the Bavarian State Ministry of Justice initiated and funded the present research project. According to our knowledge, it is the currently most comprehensive completed project on the treatment of sexual offenders in social-therapeutic prisons or units (STUs) of the justice system. In the seven STUs of Bavaria, we investigated not only (official) recidivism data, but also, for example, the type, harm and time at reoffending. In addition, the study addressed the contents of treatment, processes of assessment and selection of clients, the social climate in the units from the perspectives of staff and prisoners, problems of treatment, grant of relaxations, therapeutic aftercare in the community, specific influences of the Corona pandemic, and other issues.

Our project extended and intensified a previous study of the Criminological Service of the Bavarian Justice Ministry so that we now received data on 1,787 former prisoners from the German Central Register. We could analyze valid recidivism data of cohorts of 1,245 sexual offenders released between 2004 and 2015. As treated and untreated individuals differed in various characteristics, we applied the method of Propensity Score Matching (PSM) that included numerous variables that are relevant for treatment and outcome. Thus, we avoided selection effects and realized equivalence of the TG and CG. In a conservative approach we also used an intent-to-treat analysis that counted all individuals who dropped out after the probationary period in the TG. The samples for the recidivism analysis contained

710 inmates in the TG and 535 in the CG. The average follow-up period after release was *M* = 9.3 years (*SD* = 3.5 years). The PSM analyses showed a significant treatment effect for general reoffending (TG = 35.9 % vs. CG = 41.4 %). There were trends in the desirable direction for recidivism with a sexual offence (6.0 % vs. 8.0 %), violent offence (13.6 % vs. 14.3 %), and a "severe" offence that led to a custodial sentence of more than two years or confinement in a forensic hospital (7.5 % vs. 9.1 %). A comparison of the results for the early cohorts of the Criminological Service (2004–2008) with our expanded sample (2009–2015) showed in the four outcome categories lower reoffending rates in the TG in the more recent cohorts. Even sexual recidivism was at 4.1 % lower than before (7.3 %). The trend of lower reoffending rates in the more recent cohorts remained stable when we harmonized the follow-up period at three years for both groups. Separate PSM analyses for the more recent cohorts revealed similar results as for the older cohorts. For general recidivism, there was a statistically significant positive treatment effect. Overall, the findings of our comparison of cohorts suggest that there was an improvement of treatment practice over time. The composition of the older and more recent cohorts was similar and thus may not explain the cohort effects. However, generally more favorable conditions for rehabilitation in the society and justice system may have played a role together with treatment.

To enhance the validity of our findings we differentiated the dichotomous reoffending criteria. We developed an index of offence severity (harm index) that was independent from the previous criminal biography of the offender. It is based on the mean of the sentence range for each offence in the German Penal Code. In the TG there was a tendency of a stronger decrease of harm from the offence before imprisonment than in the CG. We also analyzed the time until the first offence after release because due to the declining age-crime curve, a delayed first recidivism may reduce reoffending in the future. This analysis revealed that the TG reoffended significantly later with regard to general offences. There was also a similar trend for violent reoffending, whereas the trend for sexual reoffending was slightly inverse. The latter may have been due to the higher proportion of offenders with child victims and paraphilias.

Although treatment in a STU is the legal norm, there was also a substantial subsample of 261 sex offenders in our TG who were not treated in a STU, but received individual or group therapy in a regular prison (e.g. because they did not agree to be transferred to an STU or there were other institutional or personal reasons). For the treatment in a regular prison the

PSM comparison with a matched equivalent CG showed positive effects: The treated individuals recidivated less often in all four criteria. Due to the smaller samples, the differences were not statistically significant. However, the pattern of findings was more homogeneous and positive than in the analysis for the STUs only. In the latter, the CG showed somewhat (but not significantly) better results than the TG in violent and serious reoffending. These findings suggest that the legal norm of sexual offender treatment in STUs may not be optimal for all cases. For an in-depth analysis it would be necessary to get more detailed information on the contents of the individual and group therapies in normal prisons, but these were not available.

In the evaluation of reoffending after treatment in a STU it was important whether the participants had completed the whole therapeutic program or dropped out earlier. As we used the methodological standard of an intent-to-treat analysis, the TG contained nearly a fifth of indviduals with earlier termination. These particpants recidivated more frequently than the CG in total and, in particular, the regular completers. With regard to general reoffending the significant difference between completers and dropouts was 35 % versus 52 %. However, non-completion had different reasons and sometimes occurred rather lately. Our detailed analyses suggest that the fact of dropout (with potential stigmatizing consequences) may not be a simple cause of enhanced recidivism, but related individual and institutional factors play a role.

With regard to the index offences, rapists had clearly higher rates than child molesters in general, violent, and severe reoffending, however somewhat lower rates in sexual reoffending. These findings applied to both treated and untreated individuals. A small group had carried out offences with both child and adult victims. ("mixed" sexual offenders). The three groups differed in various characteristics, e.g. with regard to other violent reoffences, number and kind of victims, and previous sex offences. The profile of the "mixed" offenders was similar to that of rapists and the reoffending rates were similar too. Separate PSM analyses on the recidivism of child molesters and rapists showed lower rates of treated rapists in all outcome criteria (with up to four percentage points less recidivism than the CG). Among the child molesters, the general and sexual reoffending rate in the TG was a little higher than in the CG. For both types of offenders, however, the differences between TG and CG were not statistically significant.

Important factors for offender rehabilitation are successive relaxations (e.g. opening of incarceration for visits, vacancies, or sometimes work

in the community). In concordance with the concept of STUs our data showed that two thirds of treated individuals were granted relaxations. This was only the case for about one third of the untreated sample. Individuals who got relaxations had lower recidivism rates, even when various risk variables were controlled. As aftercare is also important for rehabilitation, there recently was a substantial expansion of therapeutic aftercare in Germany, in particular through the establishment of ambulatory treatment centers. Accordingly, over three quarters of our sample received treatment orders after release. This occurred somewhat more frequently for offenders with child victims and a lower (not higher!) score in the Static-99 risk assessment instrument. There was no "booster effect", i.e. lower recidivism in treated offenders. In contrast, individuals who were not treated in prison benefitted from treatment in the community. These findings may indicate some therapeutic satiation among those individuals who received treatment in custody. Interviews with therapeutic staff suggested that therapeutic aftercare would benefit from a more structured handover to the ambulatory centers, a more flexible adaptation to individual treatment needs, and an extension of the regional treatment network.

Already our study of the early cohorts had shown substantial differences in the characteristics and recidivism rates between the seven STUs. When other variables were controlled, the outcome variance was mainly due to different risk levels in the various units. In our extended study, we addressed these differences in more detail. For this purpose, we carried out surveys with staff and prisoners on the social and therapeutic climate in the units. The internationally proven instruments EssenCES, SQL (staff), and MQPL (prisoners) contained for example subscales on the attitudes to prisoners, experience of security, support of the inmates, relation to colleagues, cohesiveness among inmates, and motivation (staff), plus well-being, personal development, experienced stress, and contact to the family (prisoners). The anonymous surveys had very satisfactory return rates of 73 % for staff and 91 % for the prisoners. In most dimensions, the mean scores were around the middle of the scale range and had plausible standard deviations. This suggests that the answers were neither sweepingly positive nor negative. The attitudes about prisoners were somewhat more positive among the therapeutic staff than in the group of prison officers. The total score in the EssenCES and positive scores in support for the prisoners and security and control in the unit correlated significantly with subjective evaluations of the effects of treatment. With coefficients up to $r = 0.62$ the relations were much stronger between the general evaluation of

the work in the STU and the total SQL score, the attitude about the unit's government, and the cohesiveness among the prisoners.

There were both similarities and differences between the mean scale values of the staff and prisoners. The most pronounced differences appeared in three units with regard to support of prisoners, whereas in other units and dimensions the scores were similar. It was not the case that one or the other unit showed generally more positive or negative scale values; the patterns were differentiated. Various scores in the climate scales correlated with the age of the inmates and with the type of the unit, i.e. whether they contained only sex offenders or also other violent offenders. In particular, experiences of drugs and exploitation, lower safety, and less cohesiveness among inmates were more frequent in STUs with a mix of sexual and other violent offenders. These findings point to similar stigmatizations that mainly child molesters experience in regular prisons.

Our extended sample also revealed highly significant differences in the recidivism rates between units, i.e. in general and violent reoffending. Here, the recidivism rates in some units were double or twice as high as in others. For sexual recidivism the heterogeneity was not significant, probably again due to the low base rates. The mean risk level in the respective inmates was a central factor for the differences between units. However, there was also a relation to various climate scales. For example, a more positive perception of staff went along with lower general recidivism in the respective units. Lower rates of violent recidivism correlated with the prisoners' views of stronger support from the staff, a more positive perception of staff, and the staff's evaluation of more cohesiveness among the inmates. These relations even persisted after statistical control of the risk level.

We also carried out in-depths interviews with members of the therapeutic staff on details of the treatment process. Although the legal regulations for STUs form the framing conditions, we found both similarities as well as substantial differences in the daily practice of the seven units. For example, there were differences in the initial assessment and selection of potential inmates, frequency of individual and group therapy sessions, use of various structured programs, granting of relaxations, contacts with family members, and the organization of ambulatory aftercare and treatment. The units adhered to the internationally prevailing Risk-Need-Responsivity (RNR) Model as a basic concept, but the concrete arrangements were rather heterogeneous. Local experiences and staff preferences seem to play a role for such differences. The evidence base for specific approaches is mostly unclear, so that it may be helpful to have a regular exchange on specific

topics and practice-oriented small-scale evaluations. For example, aspects of denial are interpreted as obstacles for a delict-oriented therapy and as a reason for refusal of a potential candidate for a STU, although national and international research suggests that there is no consistent relation to recidivism.

The interviews and parts of the questionnaires also dealt with problems in the treatment process. These were mainly attributed to the prisoners: e.g. lack of motivation to change, insufficient cooperation during therapy, disciplinary misconduct, dissemination of confidential information, not enough reflection about the own sexual offence, denial of delict aspects, conflicts between inmates, low competence for introspection and empathy, cognitive deficits, language limitations, personality disorders, disappointment about no granted relaxations and early release.

We also gathered many other data on the units, the staff, and the prisoners that would overload a brief summary. Readers who are interested in a specific topic may inspect our detailed list of contents. With regard to the prisoners, the data showed that there is a tendency to select more child molesters for treatment. In the staff survey, the vast majority was very satisfied with their work in the STU. However, both the therapeutic staff and the prison officers had the impression that their work in a STU is not properly valued from colleagues of regular prisons. Here it may be worthwhile to have some guided realistic information exchange. We also gathered various data on the STUs in Bavaria and on the influence of the Covid-19 pandemic on the daily practice. The latter showed that constraints of group therapy, the social isolation of inmates, and abolition of contacts with relatives and the world outside were particularly stressful.

In spite of our rather comprehensive data gathering and analyses, we are aware of various limitations of our study. For example, we could not start our research before the implementation of the treatment. As in many other psychosocial treatment studies, a true experiment with a randomized allocation of individuals to the TG and CG was not possible for legal and practical reasons. However, through our thorough application of PSM we could realize a sound quasi-experimental evaluation design that enabled valid data on recidivism.

It is a second limitation that we repeatedly reported and (cautiously) interpreted non-significant tendencies and trends in our findings. This is justified because we already applied a conservative strategy and used two-sided significance tests and an intent-to-treat analysis that allocated dropouts to the TG. A simple reliance on statistical significance thresholds

has been frequently criticized, even among statisticians and in prestigious journals. Mere significance reports can be misleading depending on the research theme, context, and results. Especially in the not yet consolidated field of sexual offender treatment, it is important to considerate absolute differences and consistent patterns of results. This is also justified because we analyzed a large sample and more statistical power was not possible as we already carried out a total assessment of all eligible individuals in the analyzed period.

A third limitation of our study is the time difference between the treatment of the released prisoner cohorts and the more recent gathering of data on the treatment contents and problems, the institutional climate, and other variables. Of course, we cannot definitely say how much the findings of the recent surveys and interviews apply to what had happened in the STUs five years before. This was unavoidable and is a general problem of quasi-experiments with a long follow-up. However, we compared our current data on the STUs with previous studies and found clues that there were not fundamental changes in characteristics of the structures, staff, and inmates. That in our recent cohorts the treatment was somewhat more successful than before may be a plausible learning and developmental effect that points to the validity of our study. Validity is also indicated by the fact that various results were already published in five prestigious journals with tough peer-review processes.

Another limitation may be seen in our use of official recidivism data in our outcome analyses. However, we could not carry out individual direct assessments more then years after release from prison. In addition, self-reports on sexual offending in the dark field are more problematic than on less serious offences. We could also not assess other indirect criteria of rehabilitation such as employment, family relations, and substance use after release. However, research had shown that these important psychosocial criteria are often correlated with recidivism. It must also be taken into account that policy makers and the general public are mostly interested in recidivism data and a life without crime is the only legally defined aim of treatment in prisons. We differentiated official reoffending with regard to kind, harm, time, offender type etc. and critically discussed the basic methodological problems of single act criteria in contrast to multiple act criteria in personality and developmental research.

Our study aimed for a fair and practice-oriented evaluation. Therefore, our internal final report was not only sent to the Bavarian Ministry of Justice but also to the participating treatment institutions. In a workshop

with representatives from the STUs we presented our main results and had intensive discussions on implications for future practice. The feedback contained helpful comments and differentiations. In particular, these addressed issues of treatment dropout and differences in the practice of the various institutions. Another issue that had not yet be mentioned in our interviews was the recent legal regulation for prisoners with extended custody for preventive reasons. This subgroup may lead to further quantitative and qualitative problems in the STUs. The discussion with the representatives from practice was constructive and helpful. Accordingly, we included some parts of the feedback in the present monograph. Stakeholders from practice already undertook initial steps to transfer some of our findings into future treatment. The usefulness of our research is also suggested by the fact that various results have already been published in peer-reviewed journals.

2 Forschungsstand der Behandlung von Sexualstraftätern und Themen unseres Projekts

2.1 Deliktarten, Tätertypen und Behandlungsmethoden bei Sexualdelinquenz

Sexualdelikte sind ein Problem, das in vielen Ländern besondere Aufmerksamkeit in der Bevölkerung erfährt. Dementsprechend befassen sich Kriminalpolitik, Forschung und Praxis intensiv mit dieser Thematik. Dabei geht es zum einen um angemessene Strafen, die zur generellen und spezifischen Prävention bzw. Abschreckung beitragen sollen. Zum anderen wird gefragt, wie durch Behandlungsmaßnahmen Rückfälle verhindert werden können. Trotz langjähriger Forschung und Praxis wird die Behandlung von Sexualstraftätern international kontrovers diskutiert (z.B. Ho & Ross, 2012; Ho, 2015; Koehler & Lösel, 2015; Mann et al., 2012; Marshall & Marshall, 2010; Rice & Harris, 2003; Seto et al., 2008). Ein Konsens über die Wirkung der Sexualtäterbehandlung wird durch verschiedene Faktoren erschwert:

Sexualstraftaten sind vielfältig. Darunter fallen unter anderem der sexuelle Missbrauch von Kindern, Vergewaltigung, sexuelle Nötigung, exhibitionistische Handlungen oder die Herstellung, Verbreitung und Nutzung von Kinderpornographie im Internet. Bei den einzelnen Delikten gibt es verschiedene Schweregrade und manche Täter begehen auch mehrere Arten sexueller Straftaten (z.B. Kindesmissbrauch und Vergewaltigung von Erwachsenen; Link & Lösel, 2021). Die jeweiligen Delikte können zudem sehr unterschiedlich motiviert sein (z.B. primär sexuelle Motive, Dominanz, Rache, planvoll, opportunistisch). Dementsprechend gibt es auch unterschiedliche Tätertypen, die jedoch nicht einheitlich definiert und nur partiell empirisch fundiert sind (vgl. Egg, 2014; Martínez-Catena et al., 2017; Wojcik & Fisher, 2019). Bei einem Teil der Täter liegen spezifische sexuelle Präferenzstörungen bzw. Paraphilien vor, z.B. hinsichtlich des Geschlechts und Alters der Opfer oder bestimmter Vollzugspraktiken (inklusive sadistischer Varianten). Neben der weit überwiegenden Zahl erwachsener männlicher Täter gibt es auch Jugendliche und (seltener) Frauen als Sexualtäter (vgl. Gannon & Cortoni, 2010). Die Längsschnittforschung hat gezeigt, dass sich die sexuelle Straffälligkeit in jungen Jahren meistens nicht langfristig fortsetzt (Reingle, 2012); die Mehrzahl wird nur einmal einschlägig auffällig. Entgegen einer weit verbreiteten Meinung sind Sexu-

altäter auch überwiegend keine „Spezialisten", sondern begehen oft weitere Delikte wie Diebstahl, Raub oder nicht-sexuelle Gewalttaten (vgl. Jennings, 2015; Lussier, 2005). Unterschiedlich sind die jeweils zugrundeliegenden Persönlichkeitsstrukturen, die von depressiven Symptomen und sozialem Rückzug bis zu impulsiven und psychopathischen Merkmalen reichen. Bei manchen Tätern spielen auch Intelligenzdefizite eine Rolle. Und schließlich kann im Sinne von §§20, 21 StGB bei Schuldunfähigkeit bzw. verminderter Schuldfähigkeit eine krankhafte seelische Störung, tiefgreifende Bewusstseinsstörung, Intelligenzminderung oder andere schwere seelische Störung (z.B. Persönlichkeitsstörung) vorliegen.

Die Vielfalt der Deliktarten und Täterpersönlichkeiten spiegelt sich teilweise in den unterschiedlichen Behandlungsmethoden für Sexualtäter wider (vgl. Marshall et al., 2006; kurzer Überblick bei Lösel & Schmucker, 2017). Innerhalb der Hauptformen muss jedoch im individuellen Fall seitens der Praxis differenziert werden. Grundsätzlich sind für Sexualtäter alle Behandlungsansätze geeignet, die auch bei anderen Straftätern angewandt werden (vgl. Lösel, 2012, 2022). Es folgt ein Überblick zu Therapieansätzen- und konzepten, die in der Behandlung von Sexualtätern zum Einsatz kommen:

Kognitive Verhaltenstherapie (cognitive-behavioral treatment, CBT)

Diese Ansätze sind in der Behandlung von Probanden mit Sexualdelikten international am weitesten verbreitet. Nach Marshall et al. (2006) zielt man mit CBT u.a. auf biographische Probleme (z.B. eigene Missbrauchs- und Vernachlässigungserfahrungen), Selbstwertprobleme (z.B. Minderwertigkeitsgefühle gegenüber erwachsenen Frauen), kognitive Verzerrungen (z.B. Vergewaltigungsmythen), Verantwortungsübernahme (z.B. hinsichtlich der eigenen Handlungsfähigkeit), typische Kettenreaktionen (z.B. Entscheidungsabläufe von Auslösern bis zur Tat), Bewältigungsstile (z.B. passive Vermeidung und Masturbation unter Stress), soziale Fertigkeiten (z.B. „normale" Kontaktaufnahme mit erwachsenen Intimpartnern), sexuelle Präferenzen (z.B. Details sadistischer Tendenzen), Selbst-Management und Rückfallprävention (z.B. Selbstkontrolle in Risikosituationen).

Andere Probleme, die eine kognitiv behaviorale Behandlung adressiert, sind zum Beispiel eine geringe Ärger- und Wutkontrolle, fehlende Kenntnisse über Sexualität, akute familiäre Schwierigkeiten, Kommunikationsdefizite, geringe kognitive Fertigkeiten und ggf. Substanzmissbrauch. Insge-