

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des allgemeinen Vollzugsdiensts im Jugendstrafvollzug

Joscha Hausam, Robert J. B. Lehmann, Klaus-Peter Dahle

Document type

Postprint (accepted version)

This version is available at

<https://doi.org/10.17169/refubium-37580>

Citation details

Hausam J, Lehmann RJB, Dahle KP. Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des allgemeinen Vollzugsdiensts im Jugendstrafvollzug. *Forensische Psychiatrie, Psychologie, Kriminologie*. Springer Nature. 2017;11: 163–174.
DOI: 10.1007/s11757-017-0416-5

Terms of use

All rights reserved. This document is intended solely for personal, non-commercial use.

This version of the article has been accepted for publication, after peer review (when applicable) and is subject to Springer Nature's AM terms of use, but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections.

**Zum diagnostischen Nutzen einer strukturierten Erfassung von
Verhaltensbeobachtungen des Allgemeinen Vollzugsdiensts im Jugendstrafvollzug**

Joscha Hausam¹, Robert J.B. Lehmann¹ & Klaus-Peter Dahle¹

¹ Institut für Forensische Psychiatrie, Charité Universitätsmedizin

Anmerkung der Autoren

Korrespondenz: Joscha Hausam, Institut für Forensische Psychiatrie, Charité
Universitätsmedizin, Oranienburger Str. 285, 13437 Berlin, Deutschland
joscha.hausam@charite.de

Veröffentlicht am 20. März 2017 bei Springer Nature

Hausam, J., Lehmann, R.J.B. & Dahle, KP. Diagnostischer Nutzen einer strukturierten
Erfassung von Verhaltensbeobachtungen des allgemeinen Vollzugsdiensts im
Jugendstrafvollzug. *Forens Psychiatr Psychol Kriminol* **11**, 163–174 (2017).
<https://doi.org/10.1007/s11757-017-0416-5>

Zusammenfassung

Die vorliegende Arbeit befasst sich mit der Fremdbeurteilung von Inhaftierten durch den Allgemeinen Vollzugsdienst (AVD). Verhaltensbeobachtungen durch den AVD gehören vielerorts bereits zum vollzuglichen Alltag (z.B. im Rahmen von Vollzugsplankonferenzen), erlauben sie doch potentiell wertvolle Einblicke „hinter die Kulissen“, die dem therapeutischen Personal oftmals im Verborgenen bleiben. Bisherige Forschungsbefunde aus benachbarten Disziplinen (z.B. Psychiatrie) weisen indessen darauf hin, dass solche Einschätzungen insbesondere dann valide sind, wenn sie systematisch erfolgen und strukturiert erfasst werden. Zu diesem Zweck wurde im Rahmen der Evaluation einer sozialtherapeutischen Einrichtung des Jugendstrafvollzuges ein Fragebogen mit 16 Skalen zur systematischen Einschätzung beobachtbarer Verhaltens- und Persönlichkeitsmerkmale durch Bedienstete des AVD zusammengestellt. Die Skalen wiesen in einer Stichprobe mit $n = 62$ jungen Inhaftierten der Sozialtherapeutischen Abteilung der Jugendstrafanstalt Berlin überwiegend zufriedenstellende bis gute psychometrische Eigenschaften auf. Sie korrelierten mit einer Anzahl haftbezogener Merkmale (z.B. Regelverstöße) und prognostischen Risikobeurteilungen (PCL-R und HCR-20). Regressionsanalytische Untersuchungen zur Vorhersage des Therapieerfolgs liefern weitere Belege für die Konstruktvalidität des Instruments. Schließlich konnte in einer Substichprobe ($n = 26$) von Gefangenen, die nach der Erstbeurteilung rund ein Jahr in der Einrichtung verblieben waren, gezeigt werden, dass die Einschätzungen des AVD sowohl erwünschte als auch unerwünschte Verhaltensveränderungen im Behandlungsverlauf abbilden konnten. Der potentielle diagnostische Nutzen sowie Implikationen einer systematischen Erfassung von Verhaltensbeobachtungen durch den AVD werden diskutiert.

Schlüsselwörter: ‚Fremdeinschätzung‘ ‚Fremdbeurteilung‘ ‚Sozialtherapeutische Abteilung‘ ‚Allgemeiner Vollzugsdienst‘ ‚Jugendstrafvollzug‘ ‚Evaluation‘

Diagnostic value of structured behavioral assessment by prison officers in juvenile detention

Abstract

The following study investigates behavioral assessment of prisoners by correctional officers. Behavioral observation by prison officers is an important part of the daily-routine in prisons and offender treatment programs as they allow gathering information that is normally not accessible for the therapeutic personnel. Previous research from neighboring disciplines (e.g., psychiatry) shows that such reports are valid if they are conducted in a systematic and structured way. Accordingly, for the current study we compiled a questionnaire including 16 scales for prison officers working in a juvenile offender treatment program to systematically assess behavior and personality traits. In a sample of $n = 62$ juvenile offenders the scales did show moderate to good psychometric properties. The scales correlated with objective measures of institutional behavior in prison (e.g., misconduct) as well as risk assessment instruments (i.e., PCL-R, HCR-20). Regression analysis to predict therapeutic success further attested to the construct validity of the scales. A subsample of prison officers ($n = 26$) who completed a second assessment one year later were able to assess positive and negative behavioral change made during that time. This article discusses the potential diagnostic relevance of a structured behavioral assessment of offenders by prison officers as well as further implications.

Keywords: ‘behavioral assessment’ ‘juvenile offenders treatment’ ‘SWAP’ ‘prison officers’ ‘evaluation’ ‘youth detention’ ‘behavioral observation’

Einleitung

Der folgende Beitrag befasst sich mit der Güte von Beurteilungen des Verhaltens junger Strafgefangener durch Mitarbeiterinnen und Mitarbeiter des Allgemeinen Vollzugsdienst (AVD). Dahinter steht das Anliegen, den diagnostischen Nutzen solcher Verhaltenseinschätzungen zu erkunden, beispielsweise für die Beurteilung der Wirksamkeit des Jugendstrafvollzugs. Immerhin stellt der AVD hier die mit Abstand größte Berufsgruppe und ist unmittelbar mit der Beaufsichtigung, Versorgung und Betreuung der Gefangenen beauftragt. Im Fokus des vorliegenden Beitrags steht jedoch die besondere Klientel junger Gewalt- und Sexualstraftäter, die in einer Sozialtherapeutischen Abteilung (SothA) des Jugendvollzugs untergebracht ist. Die SothA hat hier die Aufgabe, bei einer besonders gefährdeten Zielgruppe junger Täter durch die Kombination unterschiedlicher Behandlungsangebote, insbesondere aber auch durch die Schaffung eines therapeutischen Milieus auf den Wohngruppen, die ihnen anvertrauten Gefangenen zu einem Leben in sozialer Verantwortung ohne Straftaten zu befähigen. Hierfür sollen die Mitarbeiterinnen und Mitarbeiter des AVD zusätzlich zu den oben bereits genannten Aufgaben die therapeutischen Fachdienste bei der Umsetzung therapeutischer Angebote unterstützen, wofür sie oftmals besonders geschult werden. Dahinter steht die Überlegung, dass durch die Art der Umsetzung eines milieutherapeutischen Konzepts gerade der AVD einen starken Einfluss auf die Gefangenen ausübt. Nicht zuletzt verbringt er die meiste Zeit mit den Inhaftierten und beeinflusst insoweit maßgeblich den Umgang und das Klima in den Wohngruppen.

Das Konzept der therapeutischen Wohngruppe stand in Deutschland schon früh im Fokus der integrativen Sozialtherapie. Schon Rasch (1977) stellte heraus, dass in der kriminaltherapeutischen Literatur Einigkeit herrsche, dass das Milieu in der therapeutischen Gemeinschaft eine zentrale Voraussetzung für die Effizienz von Behandlung von Straftätern darstellt. Neben einem offenen und angstfreien Klima soll es Lern-, Erfahrungs- und

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

Trainingsfelder für die Klienten bereitstellen, aber auch einen Interventions- und Beobachtungsraum für die Behandler und das Vollzugspersonal schaffen, um Fortschritte, Bedürfnisse und Defizite ihrer Klientel auch außerhalb des exklusiven therapeutischen Kontakts zu explorieren und Interventionen im Alltag zu implementieren. Grundgedanke ist dabei, dass die therapeutische Gemeinschaft eben nicht nur den Rahmen für Behandlung darstellt, sondern ihrerseits Behandlungstechnik ist, und dementsprechend alle Mitglieder der Gemeinschaft aktiv am Behandlungsprozess beteiligt sind. Die tägliche Interaktion zwischen AVD und dem Gefangenen definiert somit zu großen Teilen das soziale Trainingsfeld, in dem sich dann auch etwaige Lernfortschritte im Alltag der Wohngruppe bewähren müssen. Dies bedeutet auf der anderen Seite, dass potentielle Indikatoren für den Fortschritt der Inhaftierten, wie z.B. vermehrte Reinlichkeit, Sorgfalt oder zunehmend prosoziales Verhalten, letztlich nur durch diejenigen Personen eingeschätzt werden können, die ihrerseits Teil des relevanten Beobachtungsraums sind und dadurch den Fortschritt wahrnehmen können. Hier wird die besondere Bedeutung der Beurteilungen der Klienten durch den AVD deutlich, da sie in unmittelbarem alltäglichem Kontakt mit dem Klienten steht.

Vor diesem Hintergrund diskutieren Westen und Weinberger (2004) die diagnostische Relevanz klinischer Verhaltensbeobachtungen und kommen zu dem Schluss, dass die möglichen Nachteile (z.B. Beobachter-Bias) durch die Vorteile (z.B. erfahrungsbasierte Expertise) aufgewogen werden. Insbesondere in Ergänzung zu weiteren Quellen (z.B. Selbstbeurteilungen, objektive Maße) können solche Beurteilungen demnach einen Informationsgewinn entfalten. Voraussetzung hierfür ist jedoch eine verlässliche und systematische Erfassung relevanter Verhaltensbeobachtungen und in diesem Sinne entsprechend sorgfältig konstruierte Instrumente.

Nun spielen Fremdbeurteilungsverfahren im forensischen Kontext ohnehin eine herausragende Rolle, z.B. im Rahmen standardisierter Prognoseinstrumente oder bei der

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

Einschätzung psychosozialer Entwicklungen. So gibt es für den therapeutischen Kontext z.B. Fremdbeurteilungsverfahren zur systematischen Erfassung solcher Verhaltensmuster im Alltag psychiatrischer Stationen, die mit Persönlichkeitsbesonderheiten der Patienten zusammenhängen. Beispielsweise erlaubt die *Shedler-Westen Assessment Procedure* (SWAP-200; Shedler & Westen, 1998; dt. Version von Taubner et al., 2009) dem klinischen Personal die Ausprägung solcher Verhaltensweisen einzuschätzen, die für spezifische Persönlichkeitsstörungen typisch sein sollen. Auch bei der Fremdbeurteilung durch den AVD bedient man sich im Sozialtherapeutischen Kontext oft vorgegebener Listen von Verhaltenskriterien, die jedoch oft ad hoc zusammengestellt wurden, zwischen den Anstalten divergieren und vielfach auch nur bei besonderem Bedarf (z.B. bei besonderen Vollzugsplankonferenzen mit wichtigen Weichenstellungen für den Gefangenen) herangezogen werden. Ziel solcher Beurteilungsbögen ist jedoch die Erschließung von Erfahrungen, Einschätzungen und Hintergrundwissen der Bediensteten über die Insassen für die Beurteilung von Behandlungsfortschritten und prognostischen Einschätzungen. Gerade im sozialtherapeutischen Kontext wäre indessen eine systematische Erhebung zu verschiedenen Zeitpunkten im Vollzugverlauf im Sinne einer prozessbegleitenden Diagnostik wünschenswert, um frühzeitig etwaige therapeutische Fortschritte, aber auch etwaige Stagnationen und Rückschritte, zu erkennen. Dabei ist es jedoch von besonderer Bedeutung, nicht einen „inferentiellen (klinisch-urteilenden) Gesamteindruck“ abzubilden, der hinsichtlich seiner mangelnden Validität und Unschärfe vielfach kritisiert wurde (Dawes, Faust & Meehl, 1989), sondern abgrenzbare und verhaltensnahe Beobachtungen zu sammeln und anschließend zu aggregieren.

Bei der Evaluation des Jugendstrafvollzuges gibt es einige Ansätze, Fremdbeurteilungen im Rahmen von Evaluationen als Kriterium in die Beurteilung der Zielerreichung des Vollzugs einfließen zu lassen. So wurden bei der Evaluation des hessischen

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

Jugendstrafvollzugs (Kerner et al., 2015) dem sozialen Fachdienst entsprechende Fragebögen zur Fremdbeurteilung der Inhaftierten zu Haftbeginn und zu Haftende vorgelegt. Die so erfasste Eigenständigkeit und Autonomie (z.B. ist in der Lage, sich bei schwierigen Aufgaben Hilfe zu holen), das leistungsbezogene Verhalten der Probanden (z.B. zeigt Durchhaltevermögen) und die Motivation zum Ausstieg aus einer delinquenten Karriere (z.B. setzt sich selbstkritisch mit seiner Vergangenheit auseinander) nahmen aus der Sicht der Mitarbeiterinnen und Mitarbeiter des Sozialdienstes bei diesen Befragungen über alle Probanden hinweg im Verlauf der Haft signifikant zu. Auch die Freizeitgestaltung (z.B. ist in seinen Freizeitaktivitäten beständig) und das Hygieneverhalten der Probanden (z.B. achtet auf sein Äußeres) entwickelten sich aus ihrer Sicht im Haftverlauf positiv. Hinsichtlich des Vollzugsziels einer späteren Rückfallfreiheit zeichnete sich in den Fremdeinschätzungen das beobachtete leistungsbezogene Verhalten als der stärkste Prädiktor einer gelungenen Legalbewährung im ersten Jahr nach der Haftentlassung aus. Je höher das fremdeingeschätzte Leistungsverhalten am Ende der Haft war, desto geringer fiel das Rückfallrisiko in diesem Zeitraum aus. Des Weiteren reduzierte sich das Rückfallrisiko der Probanden, je geringer die fremdeingeschätzte Impulsivität und Gewaltbereitschaft bei Konflikten mit Mitgefangenen und je höher die fremdeingeschätzte Eigenständigkeit und Autonomie am Ende der Haft ausfiel.

Zusammenfassend kann also festgehalten werden, dass systematische Fremdbeurteilungsansätze für verschiedene forensisch-relevante Professionen (z.B. Therapeuten, Pflegedienst, Sozialdienst) bereits erprobt und teilweise implementiert sind. Trotz des Bedarfs und der Relevanz für die Beurteilung therapeutischer Fortschritte und ihrer potentiell prognostischen Bedeutung steht die systematische und testtheoretisch fundierte Entwicklung eines entsprechenden Instrumentariums für die Fremdbeurteilung von Inhaftierten durch den AVD jedoch noch aus. Nicht zuletzt wäre sie auch für die Beurteilung

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

der Erreichung weiterer Vollzugsziele von Belang. So plädiert beispielsweise Suhling (2012) in Ergänzung zum allgemeinen Ziel einer erfolgreichen Legalbewährung nach Entlassung für die Beachtung auch proximaler Erfolgskriterien, um Entwicklungsfortschritte abbilden zu können. In diesem Sinne wäre systematisch zu überprüfen, ob die Teilnahme an Maßnahmen (z.B. Deliktgruppe für Gewalttäter) zu konkreten Verhaltensveränderungen führt (z.B. weniger aggressives Auftreten gegenüber Mitinhaftierten), was für eine erfolgreiche Zielerreichung sprechen würde. Die Möglichkeit reliabler und systematisch erfasster Beobachtungen solcher Verhaltensänderungen im Wohngruppenalltag durch die Bediensteten der Station wäre auch für diesen Zweck ohne Zweifel von Wert.

Das Ziel der vorliegenden Studie bestand deshalb in der systematischen Untersuchung eines solchen Instruments zur Erfassung von Verhaltens- und verhaltensnahen Persönlichkeitsmerkmalen von Gefangenen durch den AVD, das im Rahmen einer Evaluation der Berliner sozialtherapeutischen Einrichtungen entwickelt wurde. Dabei wurden drei primäre Ziele verfolgt: Zunächst wurden methodische Gütekriterien der Anwendbarkeit, Objektivität und Zuverlässigkeit des Instruments überprüft. Des Weiteren erfolgte eine Konstruktvalidierung mittels externer Kriterien des Haftverlaufs und der prognostischen Risikobeurteilung sowie mittels Vorhersage des Therapieoutcomes. Zuletzt wurde die Änderungssensitivität des Instruments im Verlauf der sozialtherapeutischen Behandlung untersucht.

Methoden

Stichprobe

Die vorliegenden Daten zu den Inhaftierten der SothA der JSA Berlin wurden im Rahmen eines umfassenderen Evaluationsprojektes der sozialtherapeutischen Einrichtungen des Landes Berlin sowie der Einrichtung zum Vollzug der Sicherungsverwahrung durch das Institut für Forensische Psychiatrie der Charité - Berlin erhoben. Aus der SothA der Jugendstrafanstalt wurden dabei $N = 63$ Inhaftierte erfasst, von denen in $n = 62$ Fällen Fremdbeurteilungen durch Bedienstete des AVD vorliegen. Bei den Beurteilern handelte es sich zumeist um die jeweils zuständigen Gruppenbetreuerinnen und –betreuer des Gefangenen. Im Mittel arbeiteten diese bereits seit $M = 57,7$ Monaten ($SD = 29,1$) in der Einrichtung und verfügten insoweit über eine mehrjährige Erfahrung mit der Arbeit in einer Wohngruppe einer sozialtherapeutischen Einrichtung des Jugendstrafvollzugs. Die Inhaftierten waren zum Zeitpunkt der Einschätzung durchschnittlich $M = 20,5$ Jahre alt ($SD = 1,8$). Sie befanden sich im Mittel seit $M = 21,5$ Monaten ($SD = 13,9$) in Haft und seit $M = 12,7$ Monaten ($SD = 10,34$) in der SothA. Die große Mehrheit von ihnen war wegen nichtsexueller Gewaltdelikte (92%) verurteilt, die übrigen fünf Gefangenen wegen sexuellen Gewalttaten (8%). Ihr Strafmaß betrug im Durchschnitt $M = 3,5$ Jahre ($SD = 1,3$; Range: 1,8 – 8 Jahre).

Zur Überprüfung der Interrater-Reliabilität wurde für eine zufällig ausgewählte Teilstichprobe eine zweite Fremdbeurteilung durch einen weiteren AVD-Bediensteten eingeholt. Zu $n = 23$ Inhaftierten liegen zwei unabhängige Fremdbeurteilungen vor, die zeitnah erhoben wurden. Zur Erfassung etwaiger Veränderungen bei den Inhaftierten wurde das Instrument nach knapp einem Jahr erneut durch Bedienstete des AVD bearbeitet, soweit sich die betroffenen Inhaftierten noch in der Einrichtung befanden. Zu $n = 27$ Inhaftierten liegt eine solche Verlaufserhebung nach durchschnittlich $M = 10,8$ Monaten ($SD = 2,1$) vor.

Erhebungsmethoden und -instrumente

Verhaltens- und Persönlichkeitseinschätzung

Das zentrale Instrument zur Erfassung der Beobachtungen des AVD über ihre betreuten Gefangenen wurde im Rahmen des Evaluationsprojektes zusammengestellt und umfasst insgesamt 94 Items. Dabei handelt es sich zunächst um 54 Items zur Verhaltenseinschätzung, die durch die Arbeitsgruppe des Evaluationsprojektes formuliert wurden. Ihr Ziel war eine möglichst auf konkrete Beobachtungen im Wohngruppenalltag gestützte Erfassung von Verhaltensweisen der Inhaftierten. Die Items wurden auf fünfstufigen Likertskalen (Das genannte Verhalten wurde bei dem Inhaftierten ... noch nie ... selten ... gelegentlich ... häufig ... sehr häufig beobachtet) erfasst und nach Maßgabe inhaltlicher Überlegungen vorab den folgenden neun Skalen zugeordnet:

1. *Konstruktives Kontaktverhalten zum AVD* (12 Items, z.B. Der Inhaftierte tritt Bediensteten zugänglich, offen und ehrlich gegenüber)
2. *Prosoziales Verhalten gegenüber Mitinhaftierten* (6 Items, z.B. Der Inhaftierte ist Mitinhaftierten gegenüber hilfsbereit)
3. *Zuverlässigkeit und Eigenverantwortung* (7 Items, z.B. Der Inhaftierte achtet selbstständig auf wichtige Termine)
4. *Freizeitstruktur und Interessen* (5 Items, z.B. Der Inhaftierte kann seine Freizeit gut strukturieren)
5. *Destruktives Kontaktverhalten zum AVD* (9 Items, z.B. Der Inhaftierte verliert die Beherrschung, wenn er eine Entscheidung nicht einsieht)
6. *Antisoziales Verhalten gegenüber Mitinhaftierten* (6 Items, z.B. Der Inhaftierte schikaniert andere)
7. *Verbal aggressives Verhalten* (2 Items, z.B. Der Inhaftierte zeigt in Konfliktsituationen mit Mitinhaftierten verbal aggressives Verhalten)

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

8. *Körperlich aggressives Verhalten* (2 Items, z.B. Der Inhaftierte wird ohne erkennbaren Anlass handgreiflich)

9. *Subkulturelles Verhalten* (5 Items, z.B. der Inhaftierte nutzt unerlaubt Mobilfunkgeräte).

Darüber hinaus wurde angenommen, dass die Skalen zwei übergeordneten Dimensionen zugeordnet werden können. Die ersten vier Skalen sollten demnach die Dimension *Positives und erwünschtes Verhalten der Inhaftierten* bilden, die übrigen fünf Skalen die Dimension *Negatives und unerwünschtes Verhalten*.

Bei den verbleibenden 40 Items zur Persönlichkeitsbeschreibung handelt es sich um eine gekürzte Auswahl von Items des *Shedler-Westen-Assessment-Procedure -200* (SWAP-200, Shedler & Westen, 1998; dt. revidierte Fassung von Taubner, Stumpe & Kächele, 2009). Das Instrument basiert ursprünglich auf einem Q-Sort Verfahren und zielt darauf ab, in klinischen Kontexten durch das klinische Personal Verhaltensweisen von Patienten zu erfassen, die für spezifische Persönlichkeitsakzentuierungen bzw. -störungen typisch sein sollen. Die 200 Items lassen sich dem Manual zufolge dabei (Shedler, 2009) zwölf Faktoren zuordnen. In der vorliegenden Untersuchung wurden hiervon Items aus den folgenden sieben Skalen aufgenommen: *Psychische Gesundheit* (10 Items), *Psychopathie*, *Feindseligkeit*, *Narzissmus*, *Emotionale Dysregulation*, *Dysphorie* und *Schizoidie* (jeweils 5 Items). Die Itemauswahl erfolgte dabei nach testtheoretischen Kriterien (Items mit den höchsten Faktorladungen) und inhaltlicher Passung für den sozialtherapeutisch-straftvollzuglichen Kontext.

Haftvariablen

Mittels strukturierter Analysen der Gefangenenpersonalakten wurden weitere Variablen aus dem Vollzugsverlauf erfasst und durch Projektmitarbeiter beurteilt. In die hier vorliegenden Auswertungen wurden hiervon die folgenden Aspekte mit jeweils dichotomer Ausprägung aufgenommen, die entsprechend ihres Nichtvorhanden- bzw. Vorhandenseins mit

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

0 und 1 kodiert wurden: *eigenständige Lockerungen* im Haftverlauf, *erneute Straftaten* im Haftverlauf, *Intensivtäterstatus* bei Haftantritt und *Art der Beendigung der SothA* (regelrecht vs. Abbruch)¹. Darüber hinaus wurden *Disziplinarmaßnahmen* erfasst, wobei getrennt jeweils diejenigen Regelverstöße gezählt wurden, die sich auf aggressives bzw. gewalttätiges Verhalten, auf Suchtmittelbesitz bzw. -konsum oder auf sonstige disziplinarisch bedeutsame Vorkommnisse (z.B. Besitz unerlaubter Gegenstände, Ruhestörung, schuldhaftes Fernbleiben von Arbeit oder Schule) bezogen. Ferner wurde das *Arbeitsverhalten* (vierstufig von ‚nie gearbeitet‘ bis ‚kontinuierlich gearbeitet‘) und die *Arbeitsleistung* (dreistufige Bewertung von ‚negativ‘ bis ‚positiv‘) erfasst.

Risikoeinschätzung

Die Beurteilung der Ausprägung psychopathischer Eigenschaften mittels der *Psychopathy-Checklist-Revised (PCL-R, Hare, 2003)* und der Risikopotentiale für erneute Gewaltdelinquenz mittels des *HCR-20 Schemas (Webster, Douglas, Eaves & Hart, 1997)* erfolgte aktuarisch gemäß Manual durch Projektmitarbeiter. Die Items beider Instrumente werden dabei dreistufig nach dem Ausmaß ihrer jeweiligen Ausprägung (0 bis 2) beurteilt und zu Summenscores verrechnet. Neben den jeweiligen Gesamtwerten werden für die PCL-R die Werte der zweifaktoriellen Lösung (affektiv-interpersoneller Faktor und deviante Lebensstilmerkmale) und für den HCR-20 die Werte der drei Skalen des Instruments (historische, klinische und Risiko-Items) berichtet.

Statistische Auswertung

Die statistischen Analysen wurden mittels SPSS Version 22 für Windows durchgeführt. Die Skalenreliabilität wurde mittels Cronbach's alpha bestimmt, wobei gemäß Cichetti (1994)

¹ Für die vorliegende Untersuchung wurde unterschieden zwischen einem erfolgreichen Therapieabschluss und einem Abbruch der Behandlung in der SothA. Hierzu sei angemerkt, dass unter Abbruch auch eine Herausnahme in den Erwachsenenvollzug und Abschiebungen subsumiert wurden.

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

Werte kleiner als .7 als nicht zufriedenstellend, Werte zwischen .70 und .90 als zufriedenstellend und Werte größer als .90 als hoch zu beurteilen sind. Die Bestimmung der Interrater-Reliabilität erfolgte mittels Intraklassenkorrelationen (ICC(3,1) Einzelmaße, justiertes Modell). Den Richtlinien von Landis und Koch (1977) entsprechend bezeichnet man eine Übereinstimmung bei Werten kleiner .20 als schlecht, zwischen 0,21 und 0,40 als ausreichend, zwischen 0,41 und 0,60 als mittelmäßig, zwischen 0,61 und 0,80 als beachtlich und oberhalb 0,81 als nahezu perfekt. Zusammenhänge mit den haftbezogenen und prognostischen Variablen wurden mittels bivariaten Korrelationsanalysen nach Pearson berechnet. Zur Vorhersage des Therapieabschlusses wurde eine schrittweise, rückwärts gerichtete logistische Regressionsanalyse durchgeführt. Da die Skalen der Verhaltens- und Persönlichkeitseinschätzung nicht durchweg normalverteilt waren, wurde zur Analyse der Veränderungssensitivität schließlich der nichtparametrische Wilcoxon-Test für abhängige Gruppenvergleiche angewandt. Mit der Formel $r = z / \sqrt{N}$ wurde bei signifikanten Unterschieden die Effektstärke berechnet, wobei nach gängiger Konvention Werte von .1 auf einen kleinen, von .3 auf einen mittleren und über .5 auf einen großen Effekt hinweisen (Cohen, 1988).

Ergebnisse

Skalenanalysen

Die deskriptiven Kennwerte der Verhaltens- und Persönlichkeitseinschätzungen auf Skalenebene können Tabelle 1 entnommen werden. Der Wilcoxon-Test ergab dabei einen signifikanten Unterschied zwischen den Mean-Scores der gebildeten positiven und negativen Gesamtskalen ($z = -5,10$, $p < .001$; $r = -.65$), demnach der AVD zum Erhebungszeitpunkt insgesamt ein höheres Ausmaß an erwünschtem als an unerwünschtem Verhalten beobachtete.

Die interne Konsistenz der Skalen (Cronbach's Alpha) sowie die Trennschärfe der Items finden sich ebenfalls in Tabelle 1. Die Skalen wiesen demnach in großer Mehrheit zufriedenstellende bis hohe Reliabilitäten auf (im Durchschnitt $M_\alpha = .85$), lediglich die Skala für schizoidie-typisches Verhalten zeigte mit seinen fünf Items einen etwas unbefriedigenden Wert. Mit Ausnahme von zwei Items wiesen die Items auch überwiegend zufriedenstellende Trennschärfen auf. Bei den Ausreißern handelte es sich einerseits um das Item *Der Inhaftierte konsumiert Alkohol* der Skala *Subkulturelles Verhalten* sowie um das Item *Der Inhaftierte schätzt und reagiert auf Humor* der Skala *Psychische Gesundheit*.

Bei zehn von 16 Skalen (63%) ergab die Parallelbeurteilung der Gefangenen durch zwei Bedienstete des AVD signifikante Übereinstimmungen in der Fremdbeurteilung der Inhaftierten (siehe Tabelle 1). Bei 7 Skalen (44%) lag sie dabei im moderaten, bei 3 Skalen (19%) im ausreichenden Bereich. Mit Ausnahme der eher wenig einmütigen Einschätzung der Fähigkeit zum strukturierten Freizeitverhalten war der Konsens bei der Beurteilung der Ausprägung positiven Verhaltens dabei insgesamt größer als bei den eher negativ konnotierten Verhaltensweisen.

Ungefähr hier Tabelle 1

Konstruktvalidität

Tabelle 2 zeigt zunächst die bivariaten Korrelationen der Verhaltensbeurteilungen des AVD mit haftbezogenen Variablen. Sieht man von substanzbezogenen disziplinarischen Verstößen (hier nur ein Zusammenhang mit dem beobachteten Ausmaß subkulturellen Verhaltens) und erneuten Delikten im Haftverlauf (schwache Zusammenhänge mit antisozialem Verhalten gegenüber Mitgefangenen und dem Ausmaß subkulturellen Verhaltens) ergaben sich vielfältige und teilweise auch beachtliche Zusammenhänge mit den Haftvariablen, die durchgängig in der zu erwartenden Richtung lagen. Auffallend war dabei, dass die Verhaltensbeurteilungen bei den zur Aufnahme in die SothA bereits älteren Gefangenen insgesamt weniger negativ ausfielen als bei den noch jugendlichen Insassen.

Tabelle 3 zeigt die korrelativen Zusammenhänge der Verhaltenseinschätzungen des AVD mit den Ausprägungen der prognostischen Risikoeinschätzungen mittels PCL-R und HCR-20. Für die PCL-R ergaben sich dabei die meisten substantiellen Zusammenhänge mit dem Faktor 1 des Instruments, der die interpersonell-affektive Seite des Psychopathiekonstrukts erfasst. Demgegenüber hing Faktor 2, der insbesondere antisozial-deviante Facetten misst, nur moderat mit beobachteten körperlich-aggressiven Verhaltensweisen und subkulturellem Verhalten zusammen. Für das HCR-20 Schema ergaben sich substantielle Zusammenhänge nur für die gegenwartsbezogenen klinischen Items der C-Skala.

Ungefähr hier Tabellen 2 und 3

Vorhersagevalidität

Für die Untersuchung der Vorhersagevalidität konnte auf eine Substichprobe von $n = 52$ Inhaftierten zurückgegriffen werden, bei denen die Behandlung in der SothA bereits

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

abgeschlossen war. Von ihnen haben 35 Inhaftierte die Behandlung nach durchschnittlich $M = 23,9$ Monaten ($SD = 7,8$) regelrecht beendet, wohingegen bei 19 Inhaftierten die Behandlung nach durchschnittlich $M = 17,7$ Monaten ($SD = 10,9$) vorzeitig beendet wurde. Bei bivariater Betrachtung ergaben sich dabei bedeutsame Zusammenhänge mit den Skalen *Konstruktiver Kontakt zum AVD* ($r = -.28$), *verbal aggressives Verhalten gegenüber Mitgefangenen* ($r = .33$), *körperlich aggressives Verhalten gegenüber Mitgefangenen* ($r = .35$), *psychopathisches Verhalten* ($r = .36$), *Feindseligkeit* ($r = .32$), *emotionale Dysregulation* ($r = .34$), *Dysphorie* ($r = .30$) und *Schizoidie* ($r = .38$) mit dem später erfolgten Therapieabbruch.

In einer schrittweise, rückwärts gerichteten logistischen Regression wurden neben den Skalen zu den Verhaltensbeurteilungen auch die Haftvariablen *Neue Straftat im Haftverlauf* und *Anzahl der Disziplinarmaßnahmen* in die Analyse mit aufgenommen, da deren große Bedeutung für einen Behandlungsabbruch in Voruntersuchungen bereits belegt werden konnte (Warth, 2015). Im Gesamtmodell der Regressionsanalyse (Tabelle 4) verblieben im letzten Schritt vier Variablen, die geschätzte 41% (Nagelkerkes pseudo- R^2) der Kriteriumsvarianz erklärten und den Therapieabschluss von 75% der Inhaftierten richtig vorhersagten.

Bemerkenswerterweise trug im Gesamtmodell das Ausmaß der durch den AVD eingeschätzten narzisstischen Verhaltensmuster negativ zur Vorhersage eines Abbruchs bei, obwohl diese Variable für sich genommen keinen bedeutsamen und zudem positiven Zusammenhang mit einem Therapieabbruch zeigte ($r = .13$). Das Ergebnis erklärt sich durch hohe Korrelationen dieser Variablen zu den durch den AVD eingeschätzten psychopathischen ($r = .81$) und feindseligen Verhaltensmustern ($r = .82$), die im Regressionsmodell in der zu erwartenden Richtung ebenfalls bedeutsam zur Vorhersage beitrugen (siehe Tabelle 4). Insofern scheint es, dass erhöhte Werte narzisstischer Verhaltenszüge dann einen positiven Prädiktor für einen erfolgreichen Therapieabschluss darstellen, wenn sie nicht gleichzeitig mit erhöhten Psychopathie- und/oder Feindseligkeitswerten einhergehen.

Änderungssensitivität

Die abschließenden Auswertungen zur Änderungssensitivität des Instruments beziehen sich auf die eingangs erwähnte Substichprobe von $n = 27$ Inhaftierten, für die Fremdeinschätzungen zu zwei Zeitpunkten vorliegen. Die Differenzwerte und die inferenzstatistischen Ergebnisse sind Tabelle 5 zu entnehmen. Der Wilcoxon-Test für abhängige Stichprobe ergab dabei signifikante Veränderungen zwischen den beiden Messzeitpunkten bei fünf der neun Skalen der Verhaltenseinschätzungen im Verlauf, die jeweils eine mittlere bis große Effektstärke ($r = -.42$ bis $-.62$) aufwiesen. Demnach nahmen zwei der positiv-erwünschten Verhaltensbereiche (prosoziales und strukturiertes Freizeitverhalten), aber auch drei der negativen Facetten (antisoziales, körperlich-aggressives und subkulturelles Verhalten) in der Einschätzung der Bediensteten im Behandlungsverlauf im Durchschnitt signifikant zu. Hinsichtlich der persönlichkeitsnahen Skalen zeigten sich im Verlauf hingegen keine überzufälligen Veränderungen.

Hier ungefähr Tabelle 5

Um zu überprüfen, ob der zunächst erwartungswidrig anmutende Befund einer signifikanten Zunahme sowohl positiver als auch negativer Verhaltensfacetten auf einen Effekt der Durchschnittsbildung der Testung und insoweit auf einen methodischen Artefakt beruht, wurden die Probanden anhand des Medians der Veränderung (Differenzwerte) auf der Gesamtskala positiven und der Gesamtskala negativen Verhaltens in je zwei Gruppen aufgeteilt, wodurch Probanden mit und ohne Positiventwicklung und Probanden mit und ohne Negativentwicklung unterscheidbar wurden. Die Gegenüberstellung der so entstandenen Gruppen ergab, dass Gefangene mit einer Verbesserung in den positiven Verhaltensskalen in 11 von 13 Fällen keine Verschlechterung der Gesamtskala negativen Verhaltens aufwiesen, umgekehrt zeigten 10 von 12 Gefangene mit ungünstiger Entwicklung der Gesamtskala negativen Verhaltens keine Verbesserung auf den positiven Verhaltensskalen ($\chi^2_{(df=1)} = 8,58$,

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

$p < .01$). Tatsächlich zeigten die Klienten mit positivem Verlauf auch im Vergleich der Ausprägung der Negativskalen vom ersten zum zweiten Messzeitpunkt keinerlei nennenswerte Veränderung ($Z = -0,12$, $p = .93$), die Klienten mit negativem Verlauf hingegen eine dezente, aber nicht signifikante Verschlechterung der Positivskalen ($Z = -1,20$, $p = .25$). Die Befunde deuten insoweit darauf hin, dass die Bediensteten des AVD in der Lage waren, mit ihren Verhaltensbeobachtungen Probanden mit positivem und Probanden mit negativem Entwicklungsverlauf zu differenzieren.

Diskussion

Mit der vorliegenden Arbeit wurde ein Instrument zur systematischen Fremdbeurteilung von Verhaltens- und Persönlichkeitsmerkmalen junger Strafgefangener durch Bedienstete des AVD vorgestellt. Auf Grundlage praktischer Vorerfahrungen und theoretischer Überlegungen wurden hierfür zunächst 54 verhaltensbasierte Items zu neun Skalen zusammengestellt, die Beobachtungen des Alltagsverhaltens der Gefangenen in unterschiedlichen Facetten erfassen. Darüber hinaus wurde zur Erfassung persönlichkeitsnaher Verhaltensmuster eine Item-Auswahl der SWAP-200 (Shedler & Westen, 1998), die persönlichkeitsstörungsspezifische Verhaltensweisen erfassen will, getroffen und in ein Fragebogenformat überführt. Zentrales Anliegen dieser Studie war eine erste Überprüfung des so entstandenen Instruments zur Fremdbeurteilung des Verhaltens junger Gefangener in den Wohngruppen der Jugendstrafanstalt Berlin und hiermit auch eine erste Erkundung der diagnostischen Potentiale der Beobachtungsmöglichkeiten des Stationspersonals.

Zunächst ist festzuhalten, dass das Instrument und seine zugrundeliegenden Skalen eine zufriedenstellende bis hohe Reliabilität aufweisen. Auch lagen die Beurteiler-Übereinstimmungen mehrheitlich im akzeptablen Bereich, bei einigen Skalen war die Übereinstimmung indessen noch unbefriedigend. Dabei muss zunächst unbeantwortet bleiben, ob dies darauf zurückzuführen ist, dass die MitarbeiterInnen des AVD unterschiedlich „nah dran“ sind oder sich möglicherweise auch hinsichtlich impliziter Normen unterscheiden (Westen & Weinberger, 2004). Erfreulich ist indessen, dass in der Beurteilung der Klientel der SothA durch den AVD die Ausprägung erwünschter Verhaltensweisen gegenüber unerwünschtem und problematischem Verhalten insgesamt überwog.

Aufgrund des neuen methodischen Zugangs wurden a priori zwar keine spezifischen Hypothesen formuliert. Dennoch scheint es angemessen zu konstatieren, dass die Ergebnisse zur Konstruktvalidierung in ihrer Gesamtschau theoretisch plausible und insofern

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD erwartungsgemäße Zusammenhänge aufwiesen. Bemerkenswert sind hierbei zunächst die Vorzeichen der differentiellen Zusammenhänge, die bei Überschreiten des festgelegten Signifikanzniveaus nahezu ausnahmslos in die zu erwartende Richtung deuten. So wiesen beispielsweise Intensivtäter in der Einschätzung der Bediensteten ein geringeres Maß an Zuverlässigkeit und Eigenverantwortung auf, imponierten indessen mit deutlich mehr antisozialen Verhaltensweisen gegenüber Bediensteten und Mitinhaftierten sowie mit subkulturellen Aktivitäten in den Wohngruppen. Demgegenüber schienen sie nicht häufiger in verbale oder körperliche Auseinandersetzungen involviert zu sein, allerdings wurden bei ihnen gehäuft Verhaltensneigungen beobachtet, die auf maladaptive Persönlichkeitszüge hindeuten. In ähnlicher Weise wurde auch der Kreis der besonders jungen Inhaftierten gegenüber den älteren Gefangenen vom AVD als deutlich auffälliger in ihrem Verhalten beschrieben, demnach sie insbesondere als verbal und körperlich aggressiver und destruktiver charakterisiert wurden, was möglicherweise mit ihrer insgesamt als geringer beschriebenen Fähigkeiten zur emotionalen Regulierung zusammenhängen könnte.

Zieht man Variablen des Vollzugsverhaltens der Inhaftierten aus anderen Quellen und Vollzugsbereichen heran, ergibt sich ebenfalls ein in sich stimmiges Bild der Zusammenhänge mit den Fremdeinschätzungen. In der vorliegenden Arbeit wurden hierzu disziplinarische Auffälligkeiten nach Art des zugrundeliegenden Fehlverhaltens getrennt erfasst. Dabei zeigte sich zunächst erwartungsgemäß, dass das beobachtete Verhalten im Wohngruppenalltag (z.B. körperliche oder verbale Aggressionen, impulsives Verhalten oder subkulturelle Aktivitäten) ausschließlich mit den zugehörigen Disziplinarmaßnahmen im Zusammenhang stand. Dies galt indessen nicht für die Sammelkategorie der sonstigen disziplinarisch relevanten Regelverstöße, die ein deutlich heterogeneres Spektrum an Regelüberschreitung beinhaltet und in Konsequenz offenbar mit einem ebenso breiten Spektrum an Problemverhaltensweisen in der Wohngruppe assoziiert war. Zusammenhänge

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

ergaben sich aber auch mit der Kontinuität und der Leistungsbereitschaft im Arbeitskontext, das mit zahlreichen Facetten des beobachteten Stationsverhaltens zusammenhing. Nicht zuletzt deuteten die Ergebnisse darauf hin, dass die Gewährung von Vollzugslockerungen offenbar nicht nur mit der Einschätzung der Zuverlässigkeit und Eigenverantwortlichkeit korreliert ist, sondern auch mit prosozialen Verhaltensmustern sowie mit Verhaltensmerkmalen einer ausgeprägten psychischen Gesundheit.

Weiterhin wurden die Zusammenhänge der Fremdeinschätzungen mit prognostischen Risikobeurteilungen mittels PCL-R und HCR-20 untersucht. Der zweite Faktor der PCL-R wies dabei signifikante Korrelationen mit der Beobachtung körperlich aggressiven und subkulturellen Verhaltens auf. Vergleichbare Befunde wurden in der Literatur vielfach bestätigt, demnach dieser Faktor eher Personen mit einer impulsiven und dissozial geprägten Lebensführung beschreibt (Hare, 2003). Der erste Faktor wies hingegen zahlreiche und teilweise recht beachtliche Zusammenhänge sowohl mit den Skalen der Verhaltensbeschreibung - und hierbei insbesondere hinsichtlich des beobachteten Sozialverhaltens - als auch mit den persönlichkeitspezifischen Skalen auf. Auch dieses Ergebnis ist schlüssig, da dieser Faktor die charakteristischen affektiven und interpersonellen Persönlichkeitsmerkmale des Konstrukts umfasst (Hare, 2003).

Für das HCR-20 wurden die Ergebnisse separat für die drei zugrundeliegenden Faktoren ausgewertet. Auffällig waren dabei zunächst die durchgängig fehlenden Zusammenhänge mit der historischen H-Skala, die Risikomerkmale aus der Vorgeschichte der Probanden erfasst, sowie dem zukunftsbezogenen R-Faktor, der sich auf den sozialen Empfangsraum und mögliche zukünftige destabilisierende Einflüsse bezieht. Demgegenüber wiesen die Fremdbeobachtungsskalen zahlreiche Zusammenhänge mit dem gegenwartsbezogenen C-Faktor auf, der aktuelle klinische Einschätzungen umfasst. Auch dieser Befund steht in Einklang mit zahlreichen bereits vorliegenden Studien (eine umfassende Übersicht findet sich

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

bei Douglas et al., 2014), die wiederholt die Validität der C-Skala des HCR-20 zur Vorhersage institutionellen Fehlverhaltens und mangelnder therapeutischer Ansprechbarkeit belegen.

Vor dem Hintergrund vielfältiger Studien, die für beide Instrumente die Vorhersagevalidität zukünftiger Rückfälligkeit belegen (aus Deutschland z.B. Dahle, 2006), stimmen die hiesigen Befunde vorsichtig optimistisch im Hinblick auf die eingangs formulierte Frage zur etwaigen prognostischen Bedeutung von Fremdbeurteilungen des Verhaltens von Gefangenen durch die Bediensteten des AVD, sie bedarf freilich noch der späteren Überprüfung. Mit Blick auf den eingangs dargestellten Evaluationsansatz von Suhling (2012) konnte in der vorliegenden Arbeit aber bereits untersucht und gezeigt werden, dass die Verhaltens- und Persönlichkeitseinschätzungen die Zielerfüllung des „Maßnahmepakets Sozialtherapie“ (S. 171, ebd.) im Sinne eines erfolgreichen Abschlusses oder eines vorzeitigen Abbruchs der Behandlung vorhersagen konnten. Neben dem bereits aus Voruntersuchungen vermuteten Einfluss des Umfangs disziplinarischer Regelverstöße auf den Ausgang der Behandlung in der SothA, umfasst das hier nun gefundene Regressionsmodell auch drei Skalen der verhaltensbasierten Persönlichkeitseinschätzung durch den AVD. Während hohe Ausprägungen der Merkmale Psychopathie und Feindseligkeit mit einer erhöhten Wahrscheinlichkeit eines Behandlungsabbruchs einhergingen, zeigte sich im Gesamtmodell ein inverser Zusammenhang mit der Skala Narzissmus. Da diese Skala für sich genommen keine bivariante Korrelation mit dem Ausgang der Behandlung aufwies, aber mit den anderen beiden Persönlichkeitsskalen korreliert war, scheinen narzisstische Verhaltensmuster aber nur dann mit einer regelhaften Beendigung der Sozialtherapie zusammenzuhängen, wenn sie nicht gleichzeitig mit psychopathischen und feindseligen Verhaltensmustern einhergehen.

Zuletzt wurde an einer Teilstichprobe die Veränderungssensitivität der Fremdeinschätzungen untersucht. Die Beschreibung der persönlichkeitspezifischen Verhaltensmerkmale änderten sich dabei nach einer rund einjährigen Behandlungsphase nicht, was auf die Stabilität der zugrundeliegenden Persönlichkeitsstile hindeuten könnte. Bei den eigentlichen Verhaltensbeobachtungen zeigten sich hingegen auf mehreren Skalen teilweise beachtliche Veränderungen zwischen den Erhebungszeitpunkten, demnach insgesamt eine Zunahme des prosozialen Verhaltens gegenüber Mitinhaftierten sowie der Freizeitstruktur und Interessen beobachtet wurden. Allerdings wurde durch den AVD im Mittel auch eine Zunahme unerwünschten Verhaltens in Form vermehrter antisozialer Verhaltensweisen sowie körperlich aggressiven und subkulturellen Verhaltens beobachtet. Nähere Analysen zeigten indessen, dass diese zunächst widersprüchlich anmutenden Ergebnisse einer Zunahme sowohl pro- als auch antisozialen Verhaltens auf dem methodischen Artefakt einer Ausmittelung unterschiedlicher Entwicklungsverläufe beruhten, da sie verschiedene Teilgruppen der Klientel betraf. Tatsächlich schienen die AVD-Bediensteten durch ihre Verlaufsbeobachtungen in der Lage zu sein, Probanden mit positiven Entwicklungen und jene mit einem negativen Verlauf zu differenzieren.

Zusammenfassend liefert die vorliegende Arbeit erste Hinweise zur Reliabilität und Validität einer systematischen Erfassung von Verhaltensbeobachtungen durch den AVD und insoweit auf ihre grundsätzlichen diagnostischen Potentiale. Sie erlauben die vorsichtig optimistische Einschätzung, mit diesem methodischen Zugang die Beobachtungsmöglichkeiten einer Berufsgruppe mit alltäglichen Kontakten zu den Gefangenen (auch) für diagnostische Zwecke zu erschließen, um hierüber weitere Erkenntnisse über prosoziales und Problemverhalten sowie über etwaige Verhaltensentwicklungen zu erhalten. Dies wäre gleichermaßen für Zwecke der Wirkforschung, aber auch für Zwecke der qualitätssichernden Verlaufsdiagnostik

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD insbesondere im Rahmen von Behandlungsmaßnahmen von Wert. Dabei ist es freilich erstrebenswert, solche Fremdbeurteilungen nicht isoliert zu betrachten, sondern unter Berücksichtigung weiterer Perspektiven (Selbsteinschätzungen, Fremdeinschätzung der Fachdienste, testpsychologische Erhebungen usw.) zu einem Gesamtbild zusammenzufügen.

Indessen ist abschließend auf einige Einschränkungen der Studie hinzuweisen. Zunächst ist hierbei die kleine Stichprobenprobengröße, insbesondere hinsichtlich der Verlaufsmessung, zu benennen. So ließen sich faktoranalytische Untersuchungen und damit einhergehende Analysen zugrundeliegender Konstrukte ergänzend zur inhaltlichen a priori Zuordnung bislang noch nicht realisieren, kleinere Zusammenhänge mangels hinreichender Gruppengrößen nicht zufallskritisch absichern und auch der abschließende Versuch einer regressionsanalytischen Vorhersage des Behandlungsausgangs wäre sicherlich anhand größerer Stichproben zu replizieren. Im Weiteren erfolgte die vorliegende Erhebung mit Beginn des Evaluationsprojektes zunächst querschnittlich, was zur Folge hatte, dass die Fremdeinschätzungen auf der Grundlage unterschiedlich fortgeschrittener Behandlungsverläufe vorgenommen wurden und die durchgeführten Korrelationsanalysen keinen Rückschluss über die Richtung der Zusammenhänge erlauben. So ist nicht auszuschließen, dass manche Einschätzungen ihrerseits durch einen besonders positiven bzw. negativen Verlauf beeinflusst wurden oder die Genauigkeit der Beurteilungen von der jeweiligen Dauer der Beobachtungszeit verzerrt wurde. Die Studie liefert jedoch gute Gründe für die Annahme, dass es sich lohnen dürfte, die Beobachtungsmöglichkeiten derjenigen Berufsgruppe, die im Justizvollzug und insbesondere auch in einer sozialtherapeutischen Einrichtung den umfänglichsten Alltagskontakt mit den Inhaftierten hat, weiter auszuloten und auch für diagnostische Zwecke zu nutzen.

Einhaltung ethischer Richtlinien

Die Autoren geben an, dass kein Interessenkonflikt besteht.

Der Beitrag beinhaltet von den Autoren im Rahmen einer Evaluation sozialtherapeutischer Institutionen durchgeführte Untersuchungen an Menschen, die vom Senat für Justiz und Verbraucherschutz beauftragt und zuvor von der behördlichen Datenschutzbeauftragten der Charité – Universitätsmedizin genehmigt wurden.

Literatur

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological assessment*, 6(4), 284.

Cohen J (1988) *Statistical power analysis for the behavioral sciences*, 2. Aufl. Erlbaum, Hillsdale, NJ.

Dahle, K. P. (2006). Strengths and limitations of actuarial prediction of criminal reoffence in a German prison sample: A comparative study of LSI-R, HCR-20 and PCL-R. *International journal of law and psychiatry*, 29(5), 431-442.

Dawes, R. M., Faust, D., & Meehl, P. E. (1989). Clinical versus actuarial judgment. *Science*, 243(4899), 1668-1674.

Douglas, Shaffer, Blanchard, Guy, Reeves & Weir (2014). HCR-20 Violence Risk Assessment Scheme: Overview and Annotated Bibliography (Current up to January, 2014). Verfügbar unter: <https://kdouglas.files.wordpress.com/2014/01/hcr-20-annotated-bibliography-version-12-january-20142.pdf> [09.02.2017]

Hare, R. D. (2003). *The psychopathy checklist–Revised*. Toronto: ON.

Kerner, H.-J., Coester, M., Eikens, A., Stelzel, K., Wagner, U., Issmer, C., & Stellmacher, J. (2015). *Evaluierung des Hessischen Jugendstrafvollzugs*. Verfügbar unter: https://justizministerium.hessen.de/sites/default/files/media/hmdjie/jugstrvo_hessen_2009_endbericht_final_2015.pdf [31.01.2017]

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 159-174.

Rasch, W. (1977). *Forensische Sozialtherapie: Erfahrungen in Düren*. Karlsruhe: CF Müller.

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

Shedler, J., & Westen, D. (1998). Refining the measurement of Axis II: A Q-sort procedure for assessing personality pathology. *Assessment*, 5, 335-355.

Shedler, J. (2009). Guide to SWAP-200 Interpretation. Unpublished Draft, Revision Date, 7-29.

Suhling, S. (2012). Evaluation der Straftäterbehandlung und der Sozialtherapie im Strafvollzug - Ansätze zur Bestimmung von Ergebnis-, Struktur- und Prozessqualität. In B. Wischka, W. Pecher & H. van den Boogaart (Hrsg.), *Behandlung von Straftätern. Sozialtherapie, Maßregelvollzug, Sicherungsverwahrung* (S. 162-232). Herbolzheim: Centaurus.

Taubner, S., Stumpe, A. & Kächele, H. (2009). Die Shedler-Westen-Assessment-Procedure (SWAP-200): Eine neue Sprache der Persönlichkeitsdiagnostik und der Messung struktureller Veränderungen? *Psychotherapeut*, 54, 27-36.

Warth, D. (2015). Unterschiede zwischen Abbrechern und Nicht-Abbrechern sozialtherapeutischer Behandlung im Jugendstrafvollzug und Gründe für den Abbruch. Unveröffentlichte Diplomarbeit, Technische Universität Dresden.

Webster, C.D., Douglas, K.S., Eaves, D. & Hart, S.D. (1997). HCR-20: Assessing risk of violence (Version 2). Mental Health Law and Policy Institute, Simon Fraser University, Vancouver.

Westen, D., & Weinberger, J. (2004). When clinical description becomes statistical prediction. *American Psychologist*, 59(7), 595.

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

Tab. 1 Deskriptive Kennwerte, Skalenreliabilität, Item-Trennschärfen und Interrater-Reliabilität der Skalen der Verhaltens- und Persönlichkeitseinschätzung

Skala (# Items)	M (SD)	Cronbach's α	Trennschärfe	ICC
<i>Verhaltensbewertungen</i>				
Gesamtskala: Positives/ erwünschtes Verhalten	2,92 (0,62)	.944	.271 - .799	.573*
Konstruktiver Kontakt zum AVD (12)	3,04 (0,70)	.916	.402 - .830	.549*
Prosoziales Verhalten ggü. Mitinhaftierten (6)	2,98 (0,68)	.785	.435 - .676	.520*
Zuverlässigkeit und Eigenverantwortung (7)	3,08 (0,84)	.912	.625 - .861	.535*
Freizeitstruktur und Interessen (5)	2,33 (0,66)	.740	.275 - .647	.287
Gesamtskala: Negatives/ unerwünschtes Verhalten	1,96 (0,70)	.944	-.090 - .832	.365*
Destruktiver Kontakt zum AVD (9)	2,16 (0,84)	.887	.358 - .861	.314
Antisoziales Verhalten ggü. Mitinhaftierten (6)	1,88 (0,82)	.879	.563 - .815	.321
Verbal aggressives Verhalten ggü. Mitinhaftierten (2)	1,98 (0,92)	.724	.573	.445*
Körperlich aggressives Verhalten ggü. Mitinhaftierten (2)	1,41 (0,76)	.850	.742	.430*
Subkulturelles Verhalten (5)	1,88 (0,72)	.718	.089 - .713	.376*
<i>Persönlichkeitseinschätzungen</i>				
Psychische Gesundheit (10)	2,70 (0,61)	.856	.166 - .719	.547*
Psychopathie (5)	2,12 (1,02)	.903	.697 - .840	.385*
Feindseligkeit (5)	2,35 (0,98)	.898	.656 - .872	.466*
Narzissmus (5)	2,30 (1,06)	.905	.748 - .773	.267
Emotionale Dysregulation (5)	2,10 (1,02)	.923	.721 - .862	.377*
Dysphorie (5)	2,02 (0,75)	.846	.548 - .812	.279
Schizoidie (5)	2,19 (0,59)	.609	.276 - .458	.220

Anmerkungen. *p <.05.

Tab. 2 Bivariate Zusammenhänge zwischen den Skalen der Verhaltens- und Persönlichkeitseinschätzung und haftbezogenen Variablen

	Alter bei Aufnahme in SothA	Intensivtätstatus	Jemals im Haftverlauf gelockert	Neue Straftaten im Haftverlauf	Disziplinar: Gewalt	Disziplinar: Substanz	Disziplinar: Sonstige	Arbeitsverhalten	Bewertung Arbeitsleistung
Verhaltensbewertungen									
Konstruktiver Kontakt zum AVD	.080	-.160	.243	-.155	-.204	-.047	-.289*	.127	.192
Prosoziales Verhalten ggü. Mitinhaftierten	-.039	.040	.364**	.041	-.145	.167	-.199	-.041	.182
Zuverlässigkeit und Eigenverantwortung	.282*	-.362**	.317*	-.109	-.248	-.203	-.475**	.357**	.241
Freizeitstruktur und Interessen	.211	-.158	.225	-.105	-.131	.008	-.190	.065	-.090
Destruktiver Kontakt zum AVD	-.252*	.375**	-.170	.153	.149	.069	.471**	-.350**	-.409**
Antisoziales Verhalten ggü. Mitinhaftierten	-.252*	.389**	.079	.286*	.180	.104	.366**	-.280*	-.252
Verbal aggressives Verhalten ggü. Mitinhaftierten	-.401**	.204	-.203	.167	.318*	.016	.365**	-.321*	-.409**
Körperlich aggressives Verhalten ggü. Mitinhaftierten	-.299*	.180	-.262*	.241	.478**	.083	.304*	-.317*	-.405**
Subkulturelles Verhalten	-.351**	.382**	.075	.278*	.068	.500**	.459**	-.301*	-.237
Persönlichkeitseinschätzung									
Psychische Gesundheit	.142	-.139	.472**	-.010	-.112	-.146	-.317*	.158	.193
Psychopathie	-.330**	.305*	-.193	.174	.233	.139	.521**	-.327*	-.342**
Feindseligkeit	-.235	.289*	-.182	.186	.239	.089	.434**	-.340**	-.334*
Narzissmus	-.176	.324*	-.031	.110	.172	-.004	.345**	-.197	-.275*
Emotionale Dysregulation	-.373**	.250	-.244	.197	.325**	-.008	.420**	-.365**	-.293*
Dysphorie	-.233	.287*	-.360**	-.020	.061	.176	.422**	-.218	-.334*
Schizoidie	-.122	.074	-.372**	.010	.317*	-.036	.189	-.126	-.279*

Anmerkungen. *p <.05 **p <.01.

Tab. 3 Bivariate Zusammenhänge zwischen den Skalen der Verhaltens- und Persönlichkeitseinschätzung und prognostischen Risikobeurteilungen

	PCL-R: Faktor 1	PCL-R: Faktor 2	PCL-R: Total	HCR-20: Historisch	HCR-20: Klinisch	HCR-20: Risiko	HCR-20: Total
<i>Verhaltensschätzungen</i>							
Konstruktiver Kontakt zum AVD	-.374**	-.232	-.386**	-.020	-.340**	-.200	-.210
Prosoziales Verhalten ggü. Mitinhaftierten	-.356**	.092	-.155	.015	-.230	-.076	-.101
Zuverlässigkeit und Eigenverantwortung	-.283*	-.175	-.279*	-.089	-.239	-.205	-.218
Freizeitstruktur und Interessen	-.274*	-.128	-.261*	.023	-.330**	-.077	-.131
Destruktiver Kontakt zum AVD	.466**	.191	.421**	.088	.390**	.131	.240
Antisoziales Verhalten ggü. Mitinhaftierten	.304*	.243	.364**	-.090	.321*	.011	.061
Verbal aggressives Verhalten ggü. Mitinhaftierten	.230	.156	.246	-.039	.245	.114	.107
Körperlich aggressives Verhalten ggü. Mitinhaftierten	.203	.280*	.311*	-.008	.316*	.137	.159
Subkulturelles Verhalten	.073	.314*	.248	.109	.230	.082	.177
<i>Persönlichkeitseinschätzungen</i>							
Psychische Gesundheit	-.207	-.133	-.203	.000	-.215	-.125	-.125
Psychopathie	.357**	.237	.373**	.106	.307*	.153	.231
Feindseligkeit	.325*	.109	.276*	.024	.334**	.089	.165
Narzissmus	.397**	.190	.378**	.029	.365**	.058	.166
Emotionale Dysregulation	.347**	.038	.240	.070	.393**	.157	.240
Dysphorie	.232	-.063	.084	.188	.161	.174	.239
Schizoidie	.304*	.030	.199	.089	.320*	.208	.248

Anmerkungen. *p <.05 **p <.01.

Diagnostischer Nutzen einer strukturierten Erfassung von Verhaltensbeobachtungen des AVD

Tab. 4 Endergebnis einer schrittweise, rückwärts gerichteten logistischen Regressionsanalyse zur Vorhersage des Therapieabschlusses

	B (SE)	Lower	Odds Ratio	Upper
<i>Constant</i>	-3,42* (1,15)			
Psychopathie	1,50* (0,71)	1,11	4,46	18,01
Feindseligkeit	1,72 (0,92)	0,93	5,57	33,56
Narzissmus	-2,26* (1,07)	0,01	0,10	0,86
Disziplinarmaßnahmen: Gewalt	0,52 (0,30)	0,94	1,69	3,05

Anmerkungen: $R^2 = .41$ (Nagelkerke). Model $\chi^2(1) = 18,11$; * $p < .05$.

Tab. 5 Veränderungsmessung der Skalen der Verhaltens- und Persönlichkeitseinschätzung im Behandlungsverlauf

	M_{Diff} (SD_{Diff})	z	r
Verhaltens einschätzungen			
Gesamtskala: Positives/ erwünschtes Verhalten	0,21 (0,65)	-1,63	
Konstruktiver Kontakt zum AVD	0,13 (0,77)	-1,08	
Prosoziales Verhalten ggü. Mitinhaftierten	0,31 (0,71)	-2,16*	-0,42
Zuverlässigkeit und Eigenverantwortung	0,16 (0,93)	-,86	
Freizeitstruktur und Interessen	0,39 (0,73)	-2,38*	-0,46
Gesamtskala: Negatives/ unerwünschtes Verhalten	0,35 (0,62)	-2,44*	-0,47
Destruktiver Kontakt zum AVD	0,24 (0,88)	-1,16	
Antisoziales Verhalten ggü. Mitinhaftierten	0,51 (0,73)	-3,20*	-0,62
Verbal aggressives Verhalten ggü. Mitinhaftierten	0,31 (0,85)	-1,60	
Körperlich aggressives Verhalten ggü. Mitinhaftierten	0,48 (0,86)	-2,48*	-0,48
Subkulturelles Verhalten	0,33 (0,57)	-2,63*	-0,51
Persönlichkeitseinschätzungen			
Psychische Gesundheit	0,23 (0,67)	-1,70	
Psychopathy Psychopathie	0,23 (0,78)	-1,46	
Feindseligkeit	0,10 (0,90)	-,56	
Narzissmus	0,30 (0,93)	-1,52	
Emotionale Dysregulation	0,12 (0,94)	-,85	
Dysphorie	-0,08 (0,04)	-,47	
Schizoidie	-0,11 (0,72)	-,52	

Anmerkungen. *p <.05